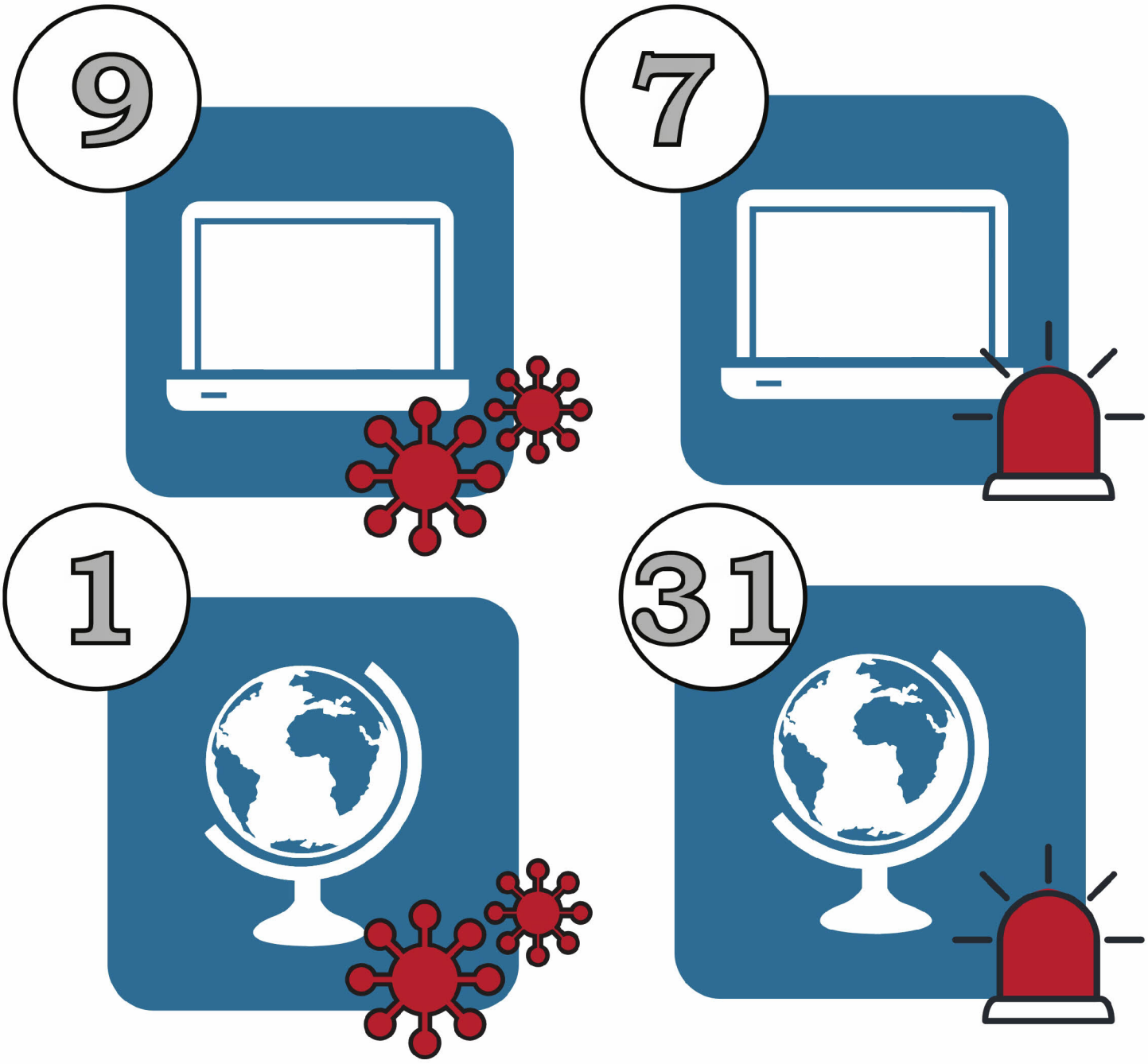
APPENDIX I.
MODELING & ANALYTICS PRODUCTS

Year 3 Semi-Annual M&E
1.1b Maps & Models



Infographic: Number of viral (left) and risk (right) characterization models (top) and maps (bottom) developed or refined between 10/1/16 – 3/31/17.

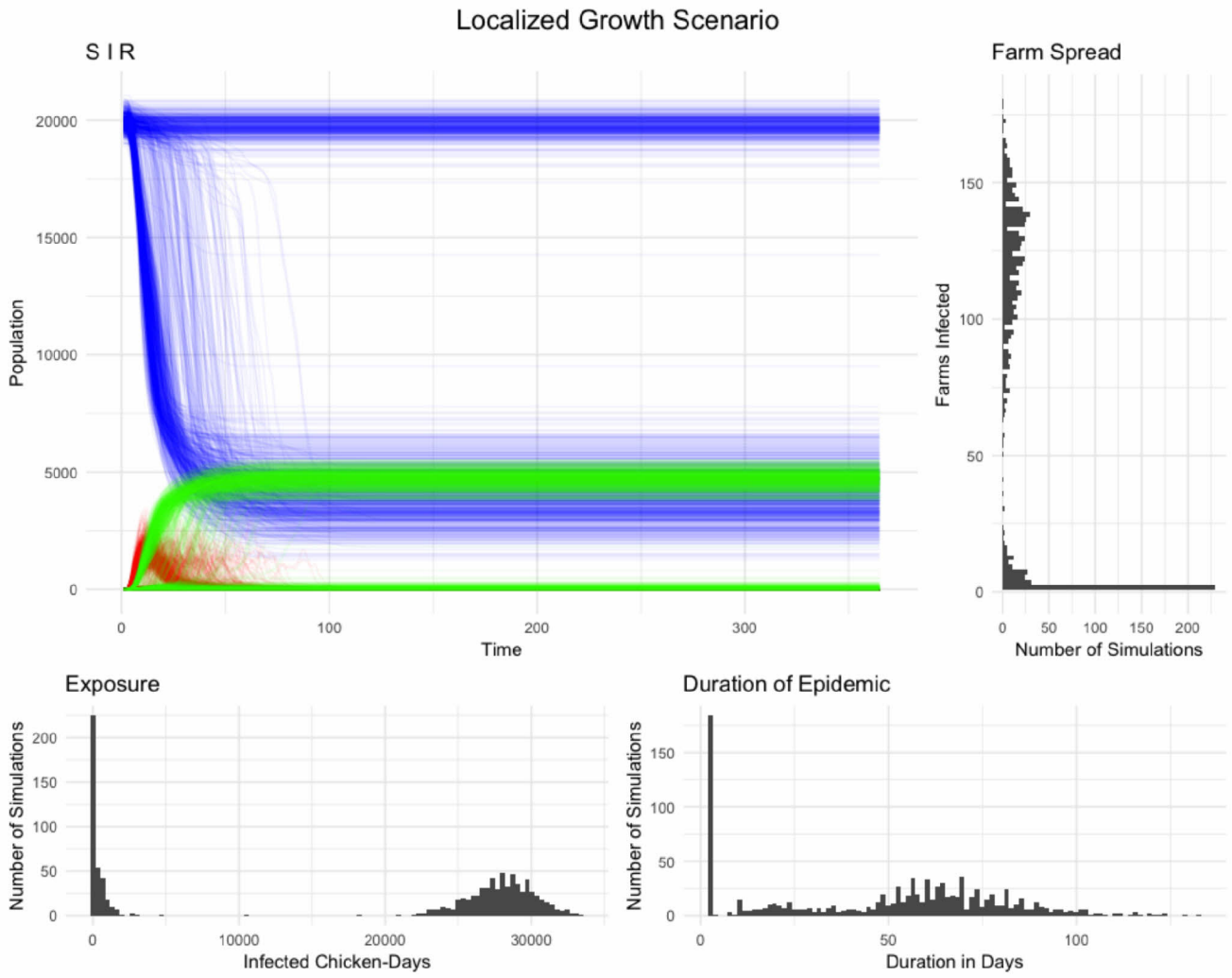
USAID
FROM THE AMERICAN PEOPLE

Emerging Pandemic Threats Program

UCDUSR0020018

1. Global

ASL 2050 – Simulating Avian Influenza Outbreaks



1. Worked with our EPT partners (FAO, P&R) on the Africa Sustainable Livestock (ASL) 2050 project to analyze how intensification of poultry production and different surveillance and control policies (e.g. culling) will impact avian influenza risk.

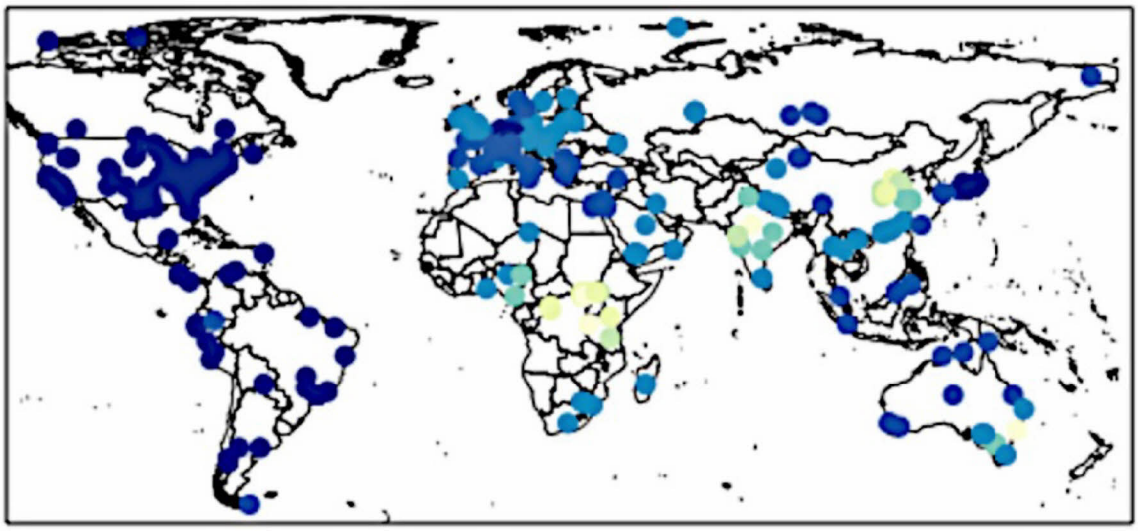
Figure shows a how a single infection of avian flu spread through a network of 200 simulated poultry farms (meta-population model, run 1000 times). SIR graph (in color) is an overlay of the susceptible, infectious, and recovered populations in each simulation; the other graphs show the distribution of epidemic duration, infectious exposure, and epidemic spread to other farms in the network. Other simulations (not shown) illustrated how different culling policies are able to reduce epidemic spread.

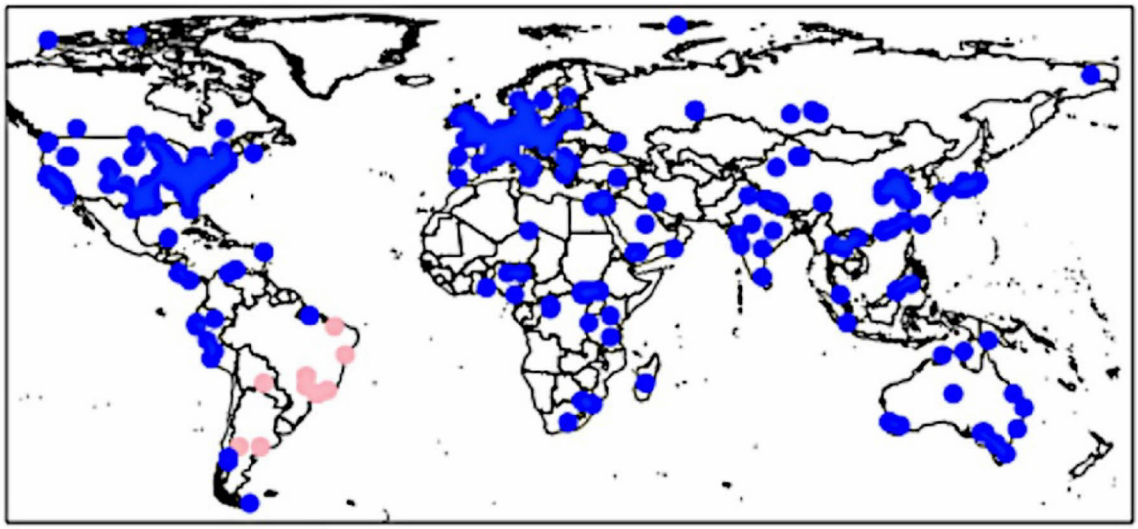
USAID FROM THE AMERICAN PEOPLE

Emerging Pandemic Threats Program

2. Global

ASL 2050 Hotspot Analysis





[97.73,98.05]
(98.05,98.38]
(98.38,98.7]
(98.7,99.03]
(99.03,99.35]
(99.35,99.67]
(99.67,100]

2. Worked with EPT-2 partners FAO to develop a hotspots analysis for the ASL 2050 project. This allows us to identify how proposed changes to livestock production in Africa will affect the risk of zoonotic disease emergence. Analyses show important regional differences in the drivers of EIDs that can be used to help policy goals.

Top figure shows emerging disease hotpots color coded by how well our model predicts the underlying factors of emergence, with darker colors having better predictive power. As an example, the bottom figure shows the relative importance of two variables, human population vs. poultry density in predicting EID hotspots. The pink dots show where poultry density is more important than human populations density.

USAID
FROM THE AMERICAN PEOPLE

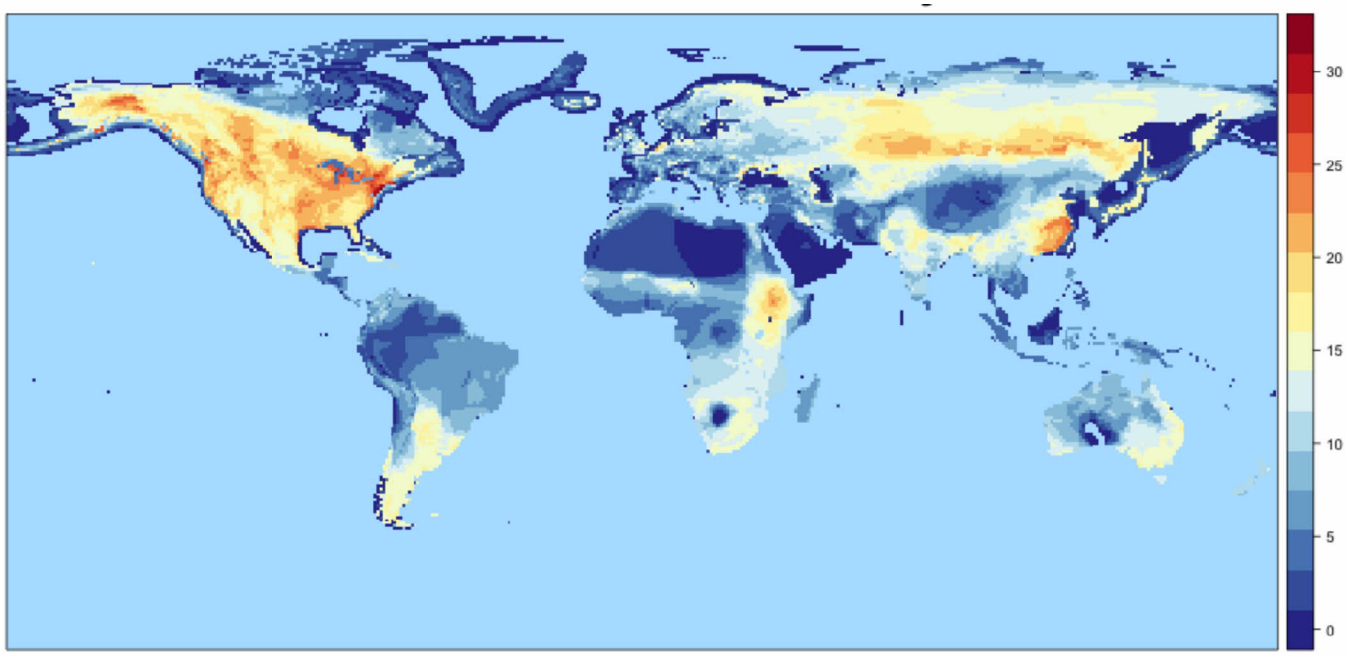
Emerging Pandemic Threats Program

UCDUSR0020020

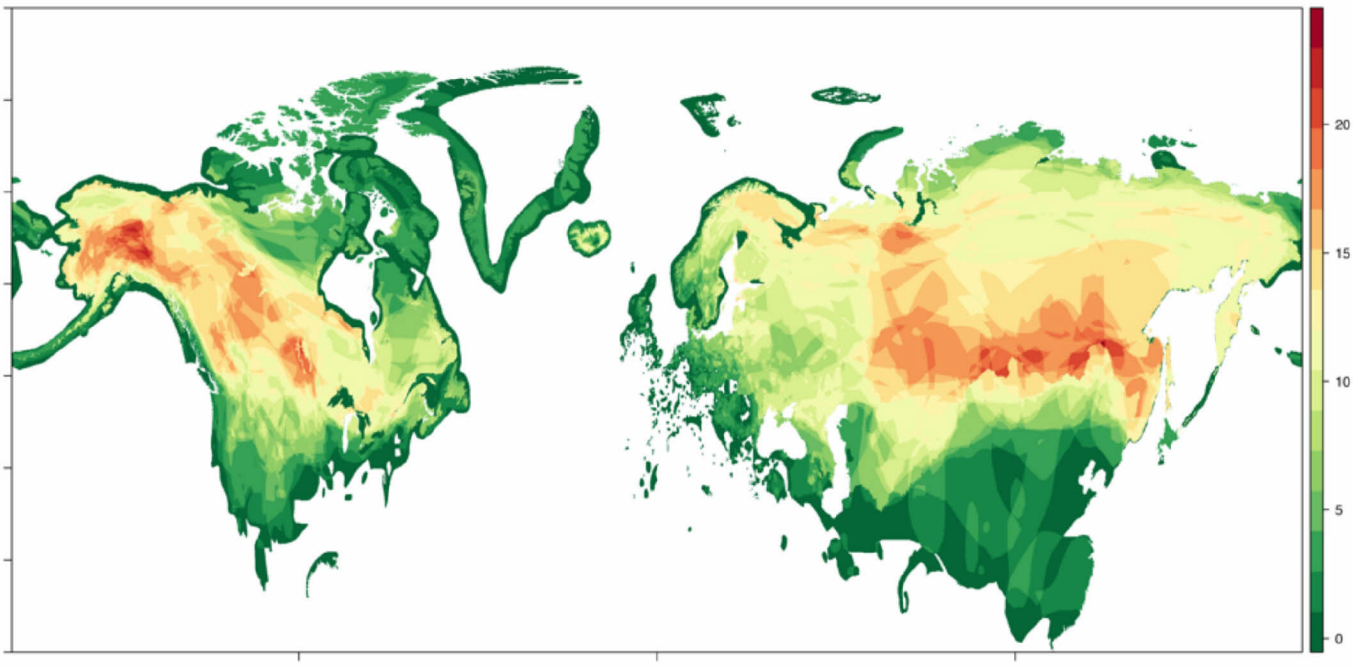
3. Global

Waterbird Diversity Maps

a.



b.



3. To better understand the global distribution of avian influenza risk and diversity, the distribution of 871 waterbird species (Anatidae) was mapped. Separate maps were created for a) non-breeding and b) breeding seasonal distribution. This map is being used to optimize sites for sampling and testing of influenza.

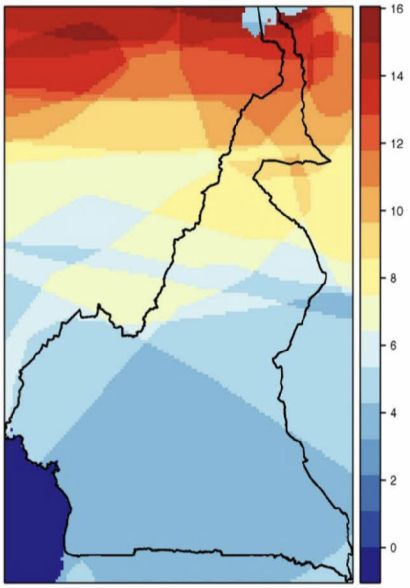
USAID
FROM THE AMERICAN PEOPLE

Emerging Pandemic Threats Program

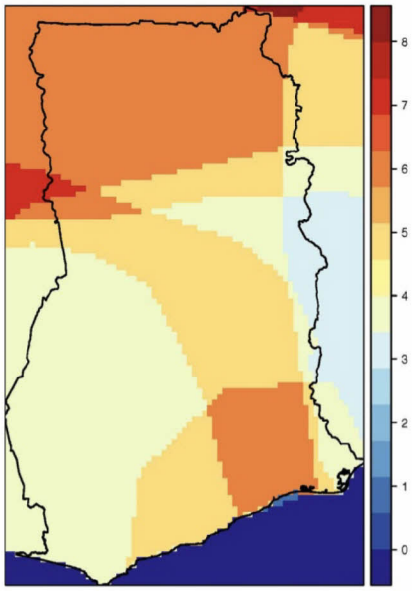
UCDUSR0020021

4. West Africa

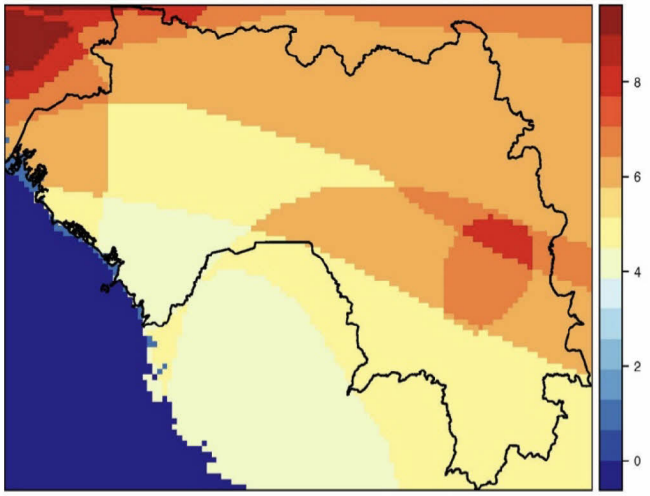
Waterbird Diversity Country-Specific Maps



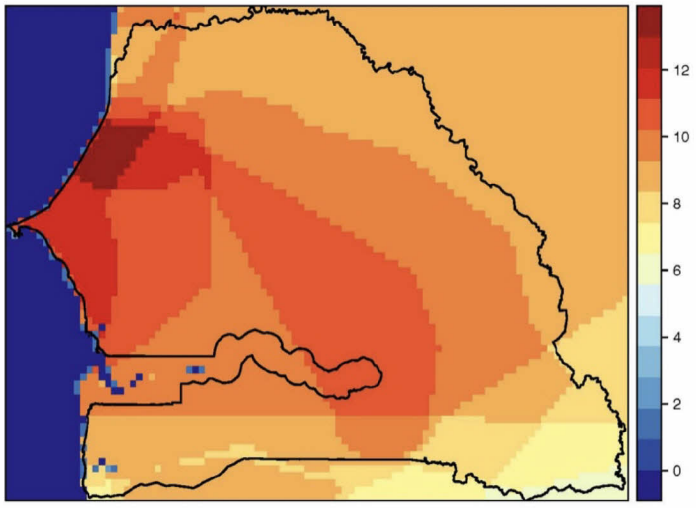
Cameroon



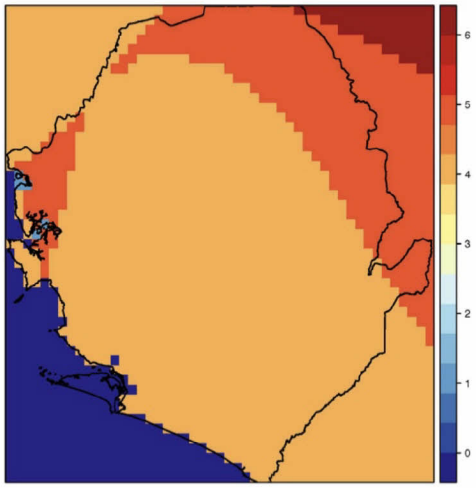
Ghana



Guinea



Senegal



Sierra Leone

4. Country-level rescaling of global waterbird species (Anatidae) distribution maps for West Africa, based data for 871 species. These maps are being used to optimize sites for sampling and testing of influenza.

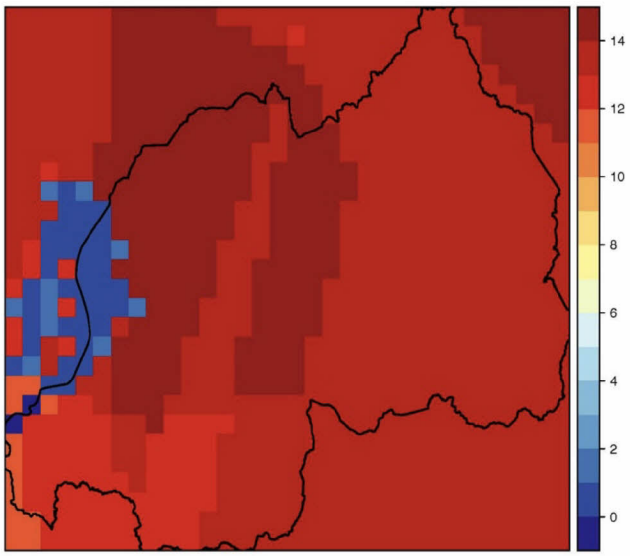
USAID FROM THE AMERICAN PEOPLE

Emerging Pandemic Threats Program

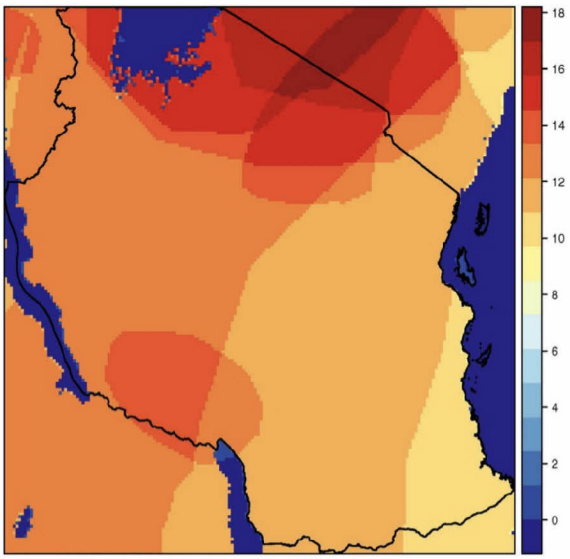
UCDUSR0020022

5. Central & East Africa

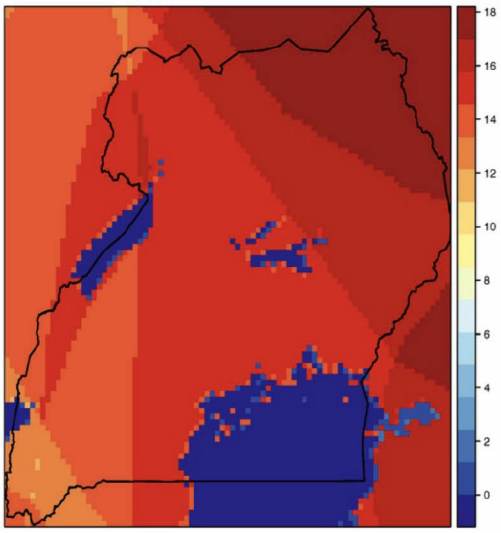
Waterbird Diversity Country-Specific Maps



Rwanda



Tanzania



Uganda

5. Country-level rescaling of global waterbird (Anatidae) species distribution maps for Central and East Africa, based data for 871 species. These maps are being used to optimize sites for sampling and testing of influenza.

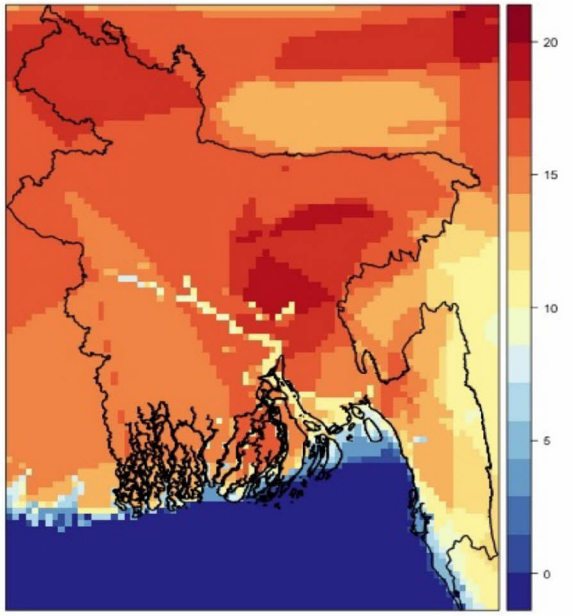
USAID
FROM THE AMERICAN PEOPLE

Emerging Pandemic Threats Program

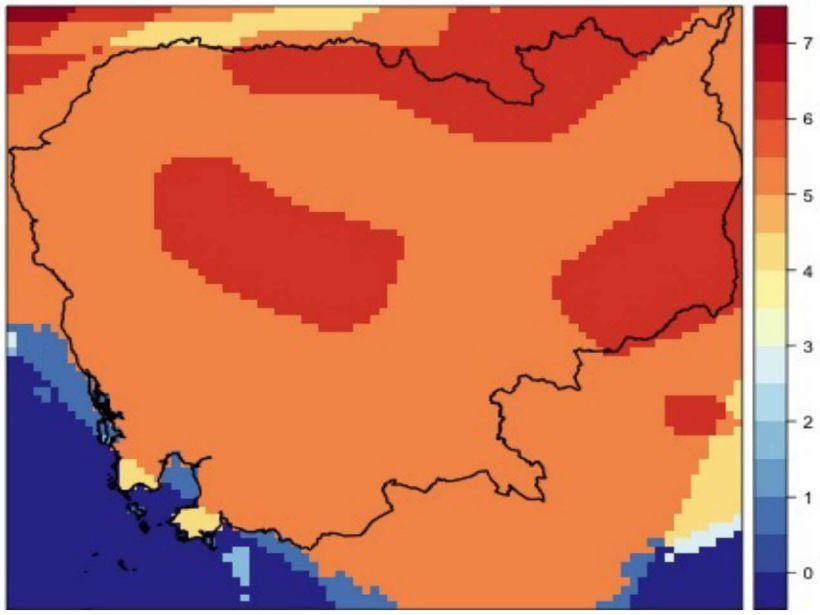
APPENDIX I.
MODELING & ANALYTICS PRODUCTS

6. Asia

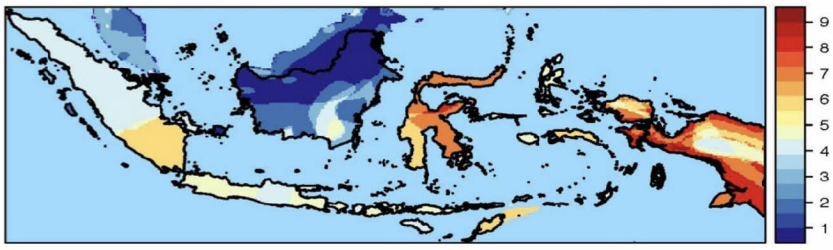
Waterbird Diversity Country-Specific Maps



Bangladesh

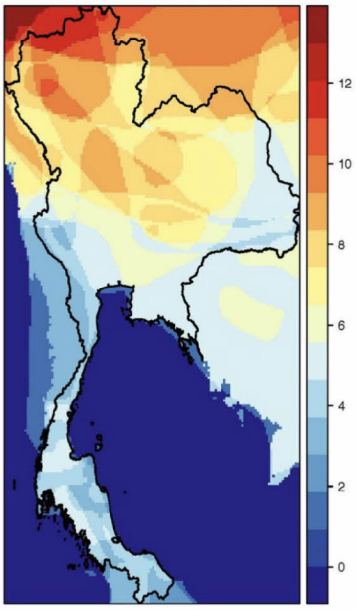


Cambodia



Indonesia

Malaysia



Thailand

6. Country-level rescaling of global waterbird (Anatidae) species distribution maps for Asia, based data for 871 species. These maps are being used to optimize sites for sampling and testing of influenza.

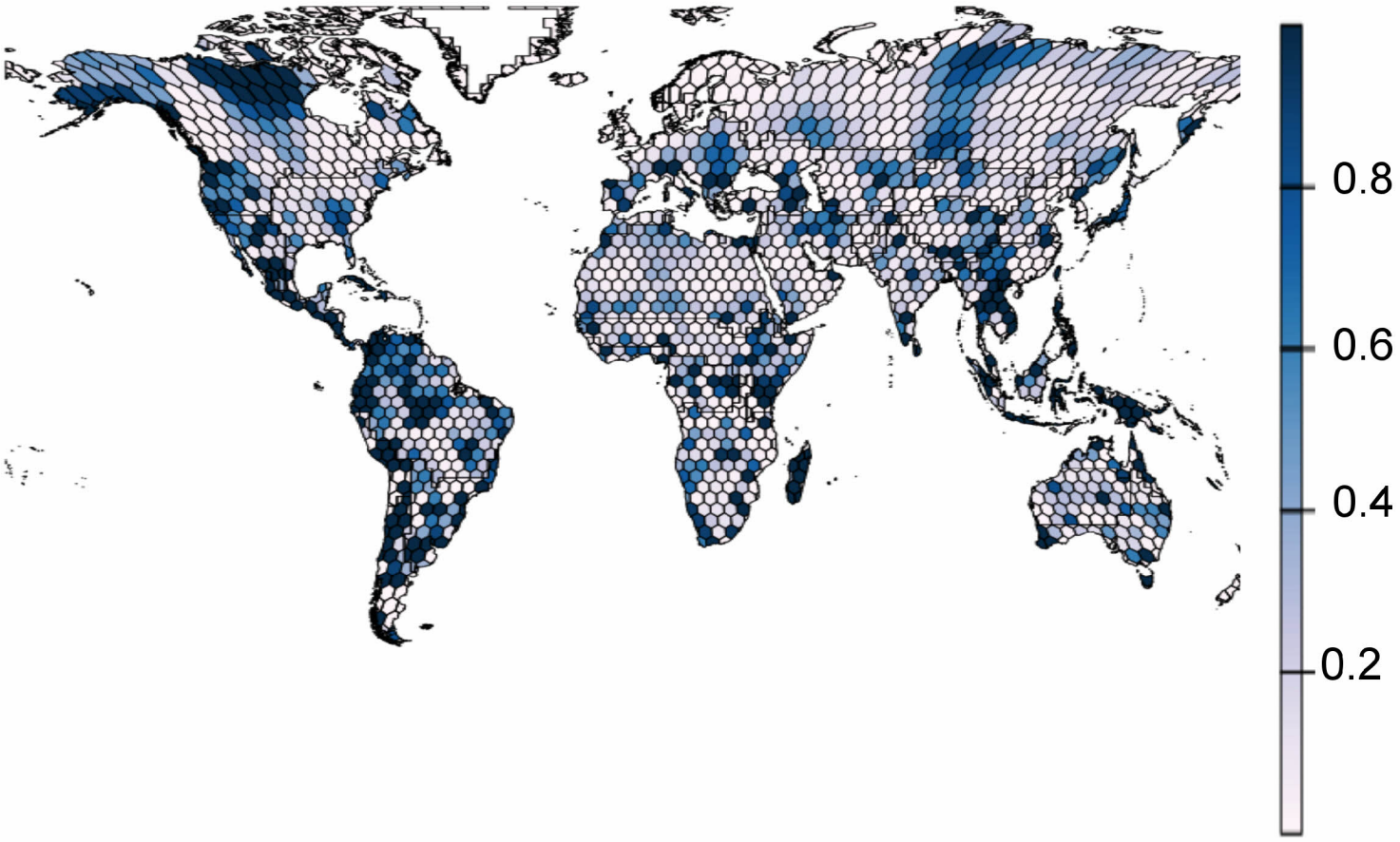
USAID FROM THE AMERICAN PEOPLE

Emerging Pandemic Threats Program

UCDUSR0020025

7. Global

Viral discovery Site Optimization Maps



7. Developed modeling approach to identify sites for optimal targeting of global viral discovery from mammal and water bird populations. This flexible solution can be used nationally, regionally or globally, and identifies the most cost-effective allocation of resources to discover new viruses from wildlife. Individual models and maps were developed for 30 countries, 10 regions, and globally.

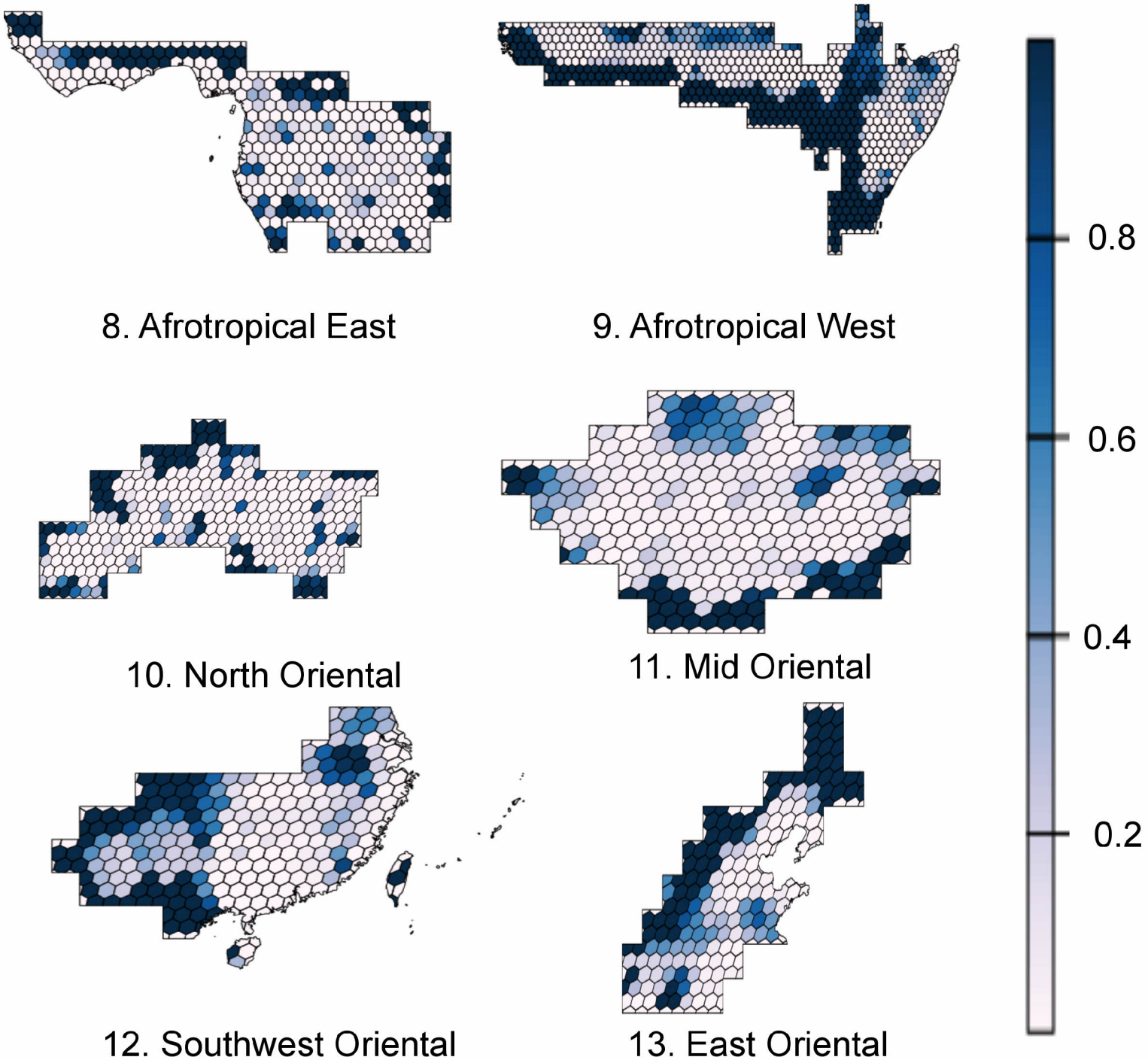
Scale bar and colors represents frequency the unit was selected in 100 model runs. Figures 8-31 show outputs for regionally optimized models.

USAID
FROM THE AMERICAN PEOPLE

Emerging Pandemic Threats Program

8-13. Regional Models

Viral discovery Site Optimization Maps



8. Afrotropical East

9. Afrotropical West

10. North Oriental

11. Mid Oriental

12. Southwest Oriental

13. East Oriental

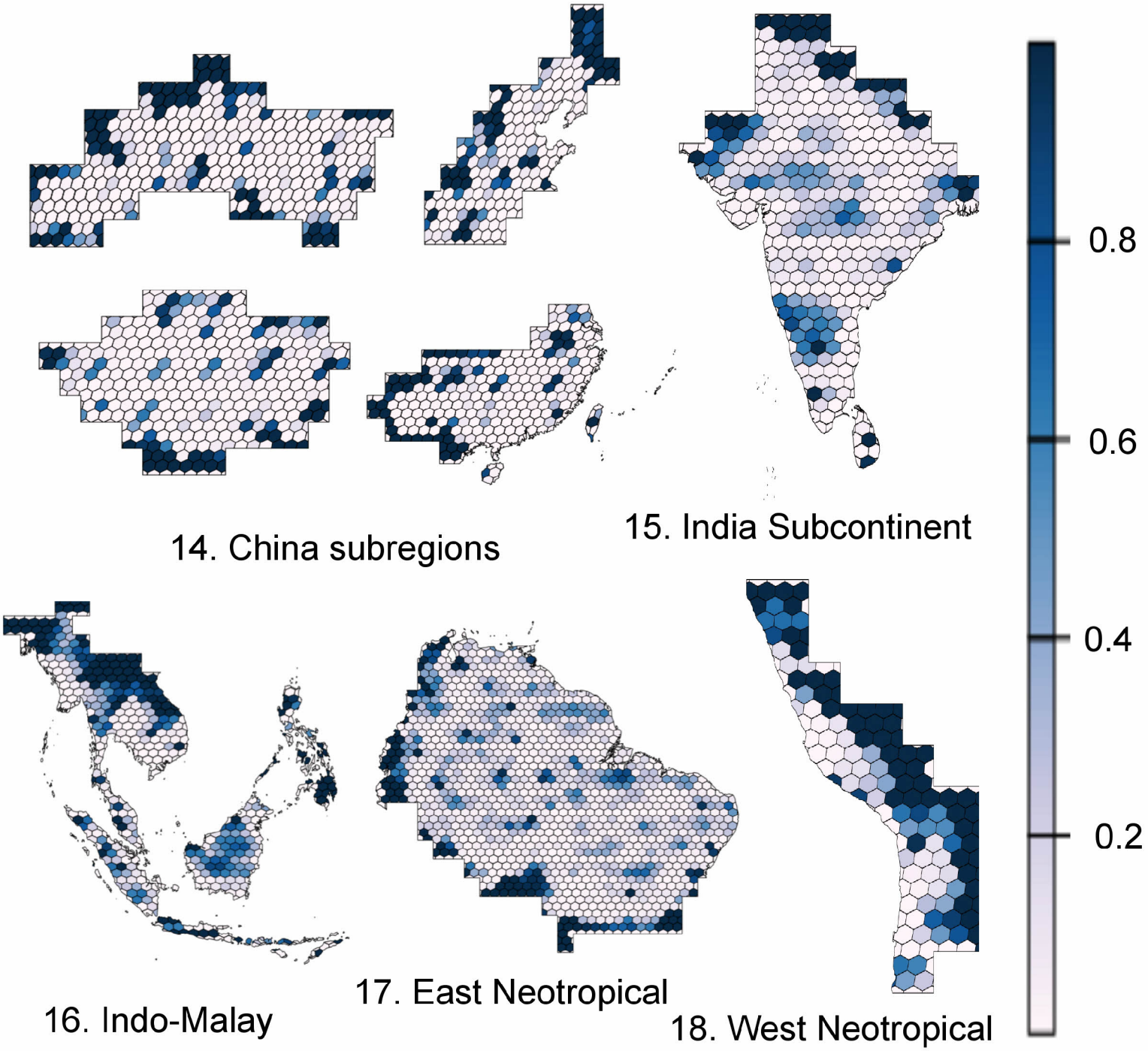
8-13. Developed modeling approach to identify sites for optimal targeting of global viral discovery from mammal and water bird populations. Six regional models shown (see Figure legend for model 7 for more details).

USAID FROM THE AMERICAN PEOPLE

Emerging Pandemic Threats Program

14-18. Regional Models

Viral discovery Site Optimization Maps



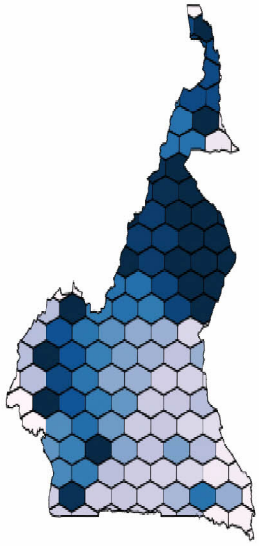
14-18. Developed modeling approach to identify sites for optimal targeting of global viral discovery from mammal and water bird populations. Five regional models shown (see Figure legend for model 7 for more details).

USAID FROM THE AMERICAN PEOPLE

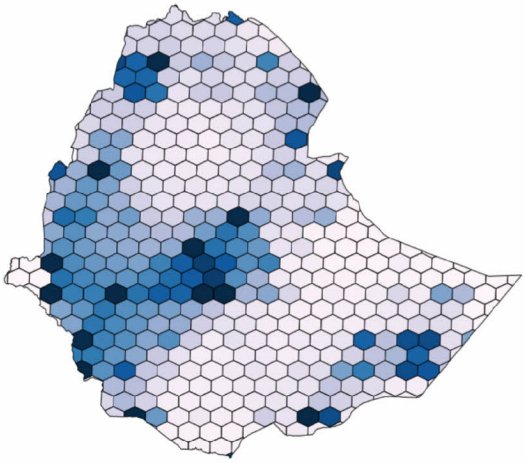
Emerging Pandemic Threats Program

19-23. Africa & Middle East

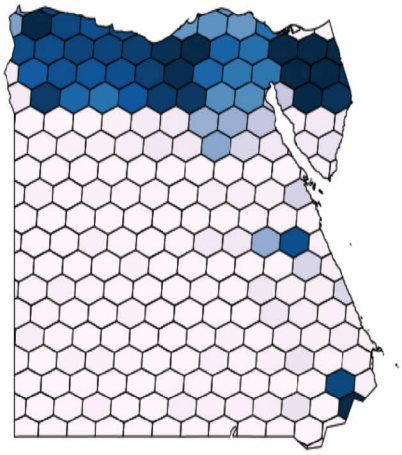
Viral discovery Site Optimization Maps



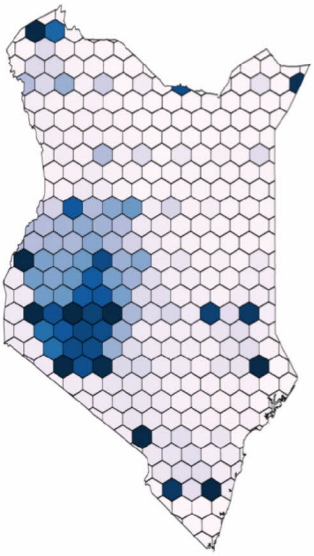
19. Cameroon



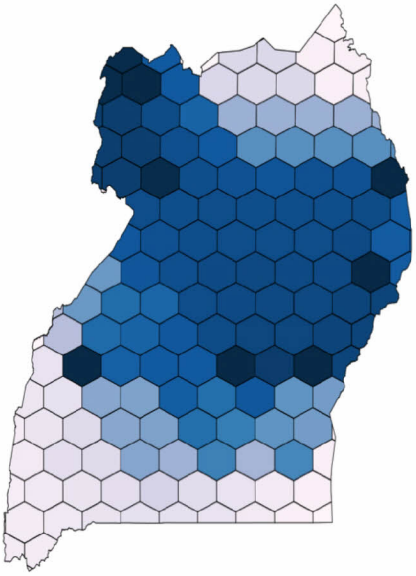
20. Ethiopia



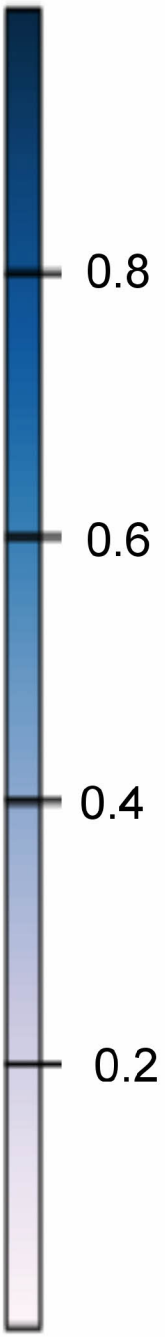
21. Egypt



22. Kenya



23. Uganda



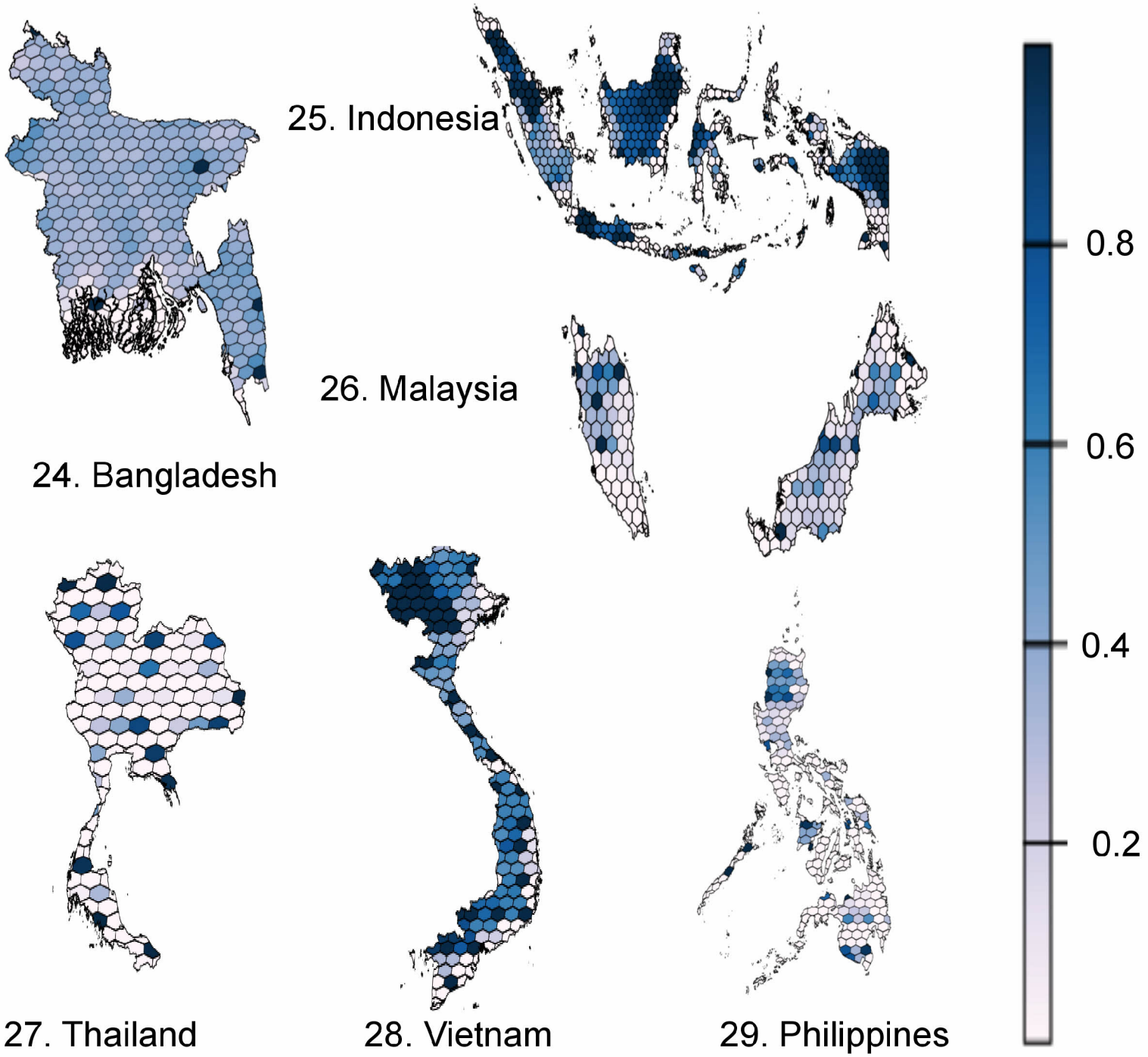
19-23. Developed modeling approach to identify sites for optimal targeting of global viral discovery from mammal and water bird populations. Five country-specific models shown (see Figure legend for model 7 for more details).

USAID FROM THE AMERICAN PEOPLE

Emerging Pandemic Threats Program

24-29. South and Southeast Asia

Viral discovery Site Optimization Maps



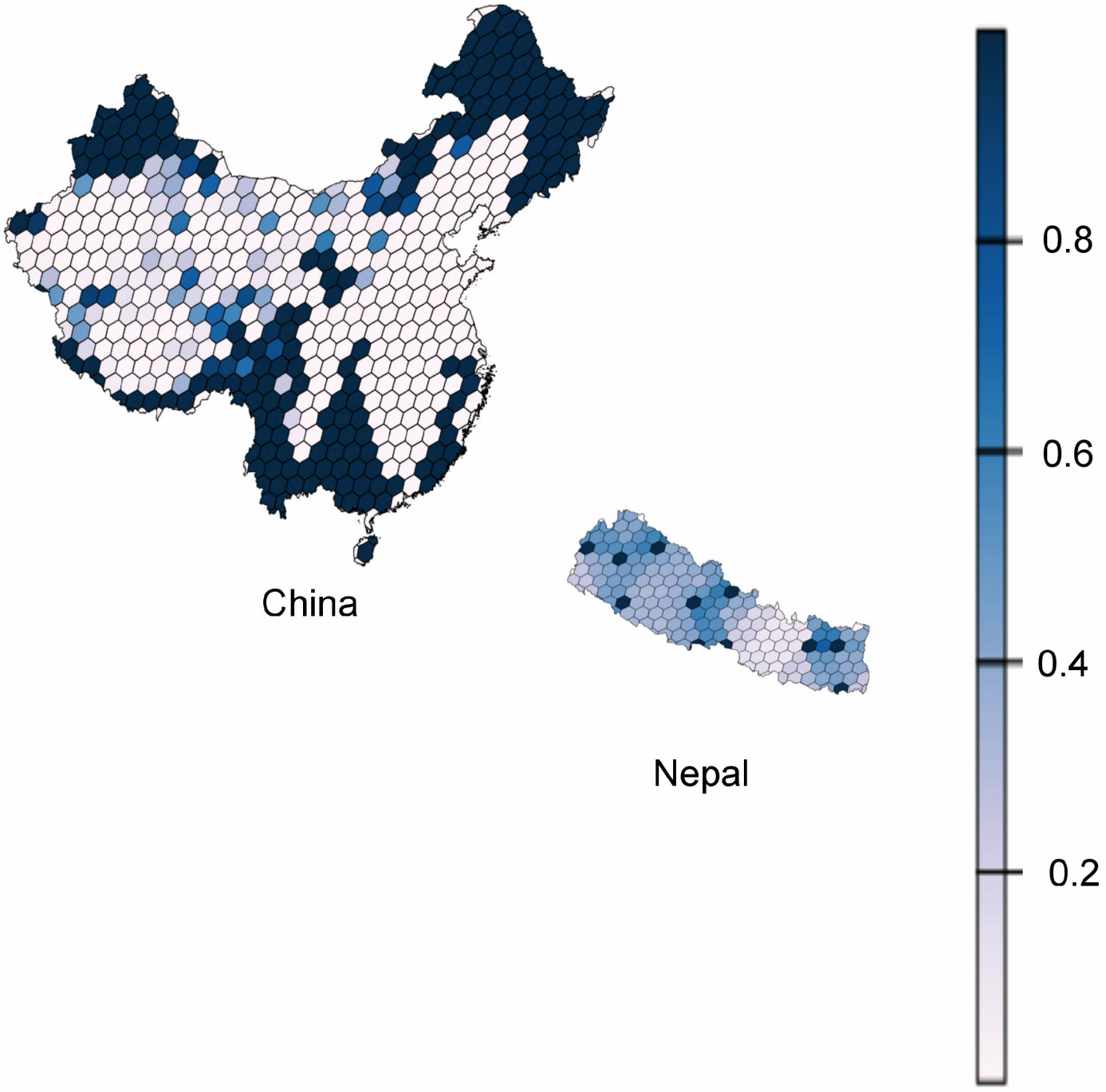
24-29. Developed modeling approach to identify sites for optimal targeting of global viral discovery from mammal and water bird populations. Six country-specific models shown (see Figure legend for model 7 for more details).

USAID FROM THE AMERICAN PEOPLE

Emerging Pandemic Threats Program

30-31. Asia

Viral discovery Site Optimization Maps



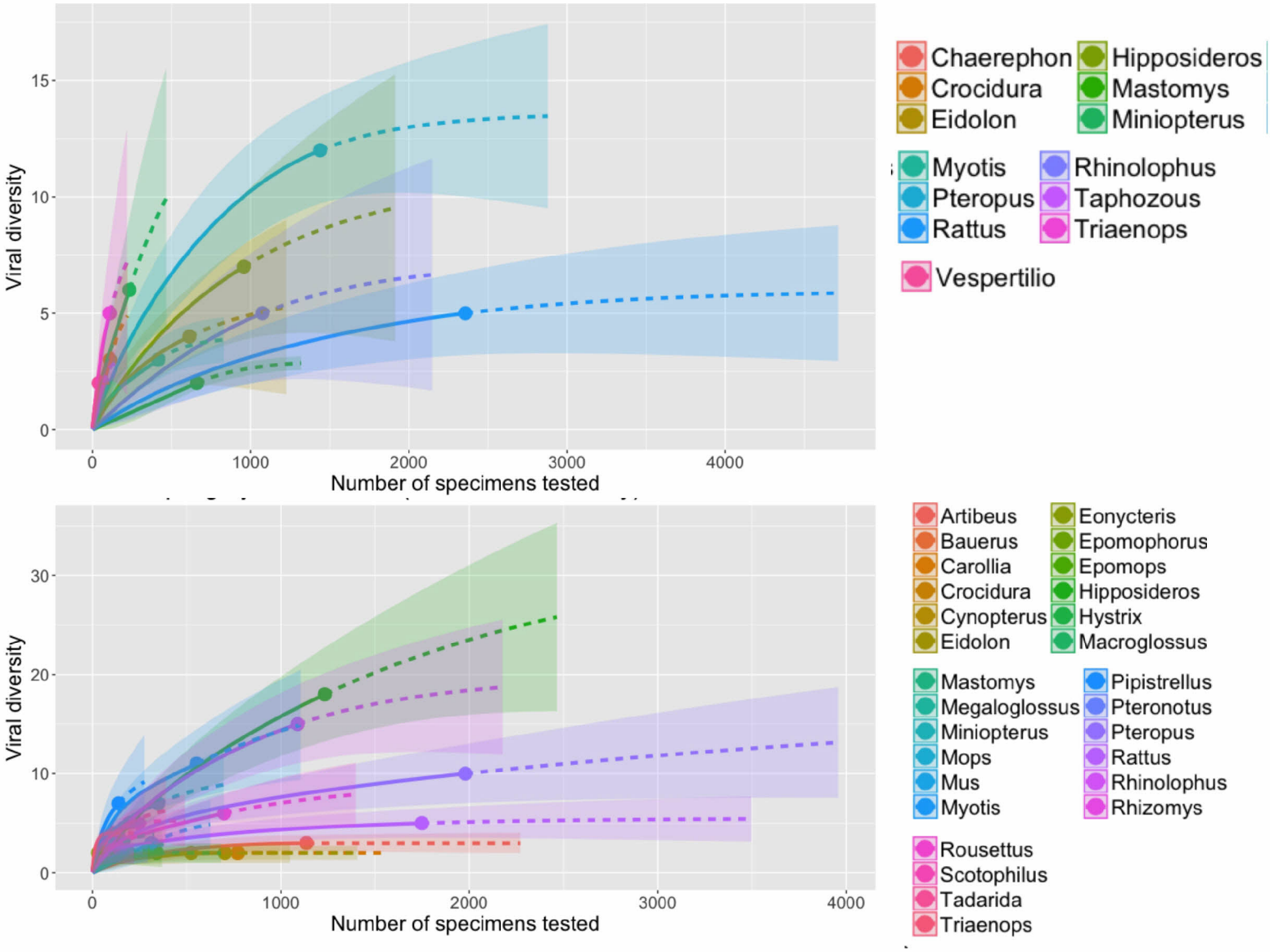
30-31. Developed modeling approach to identify sites for optimal targeting of global viral discovery from mammal and water bird populations. Two additional country-specific models shown (see Figure legend for model 7 for more details).

USAID FROM THE AMERICAN PEOPLE

Emerging Pandemic Threats Program

32. Global

PREDICT 1 Viral Accumulation Curves



32. Using PREDICT-1 data, developed an interactive *viral accumulation curve* tool and searchable table to allow in-country staff to explore viral discovery data, and compare the efficacy of different sampling protocols. Examples show viral family-specific discovery curves for a) Paramyxoviruses; and b) Coronaviruses at the level of host genus.

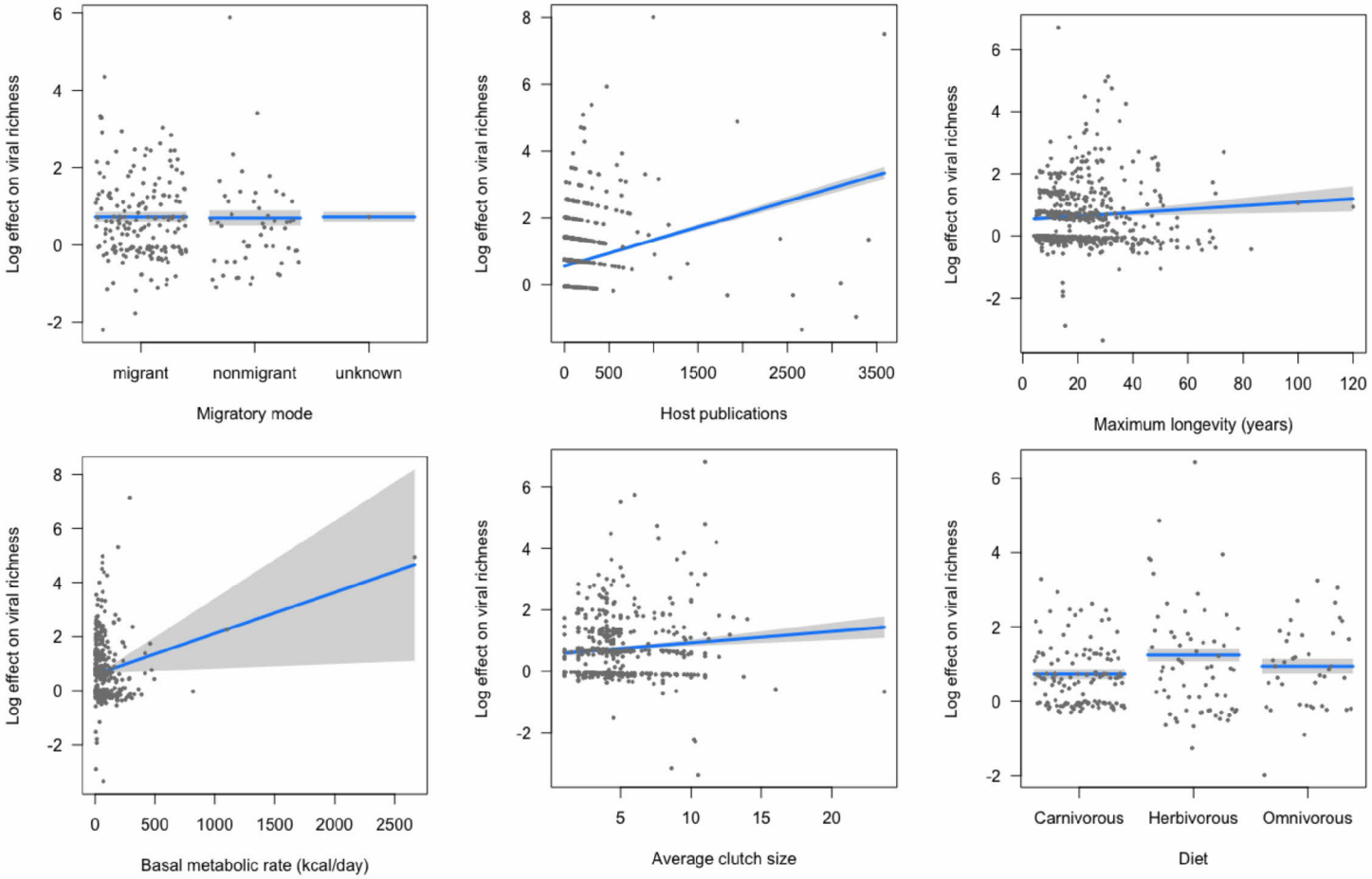
USAID
FROM THE AMERICAN PEOPLE

Emerging Pandemic Threats Program

33. Global

Avian Viral Richness Model

Preliminary Analysis from Literature



33. To better understand the diversity of viruses in birds, the M&A team analyzed how various well-known life-history traits of birds (x-axes) affects the diversity of viruses they harbor ('viral richness', y-axis)

Model uses data from 1907 avian-virus associations in the literature for 929 bird species and 142 viruses. Next steps will be to use these relationships to map the distribution of avian viral diversity globally.

USAID FROM THE AMERICAN PEOPLE

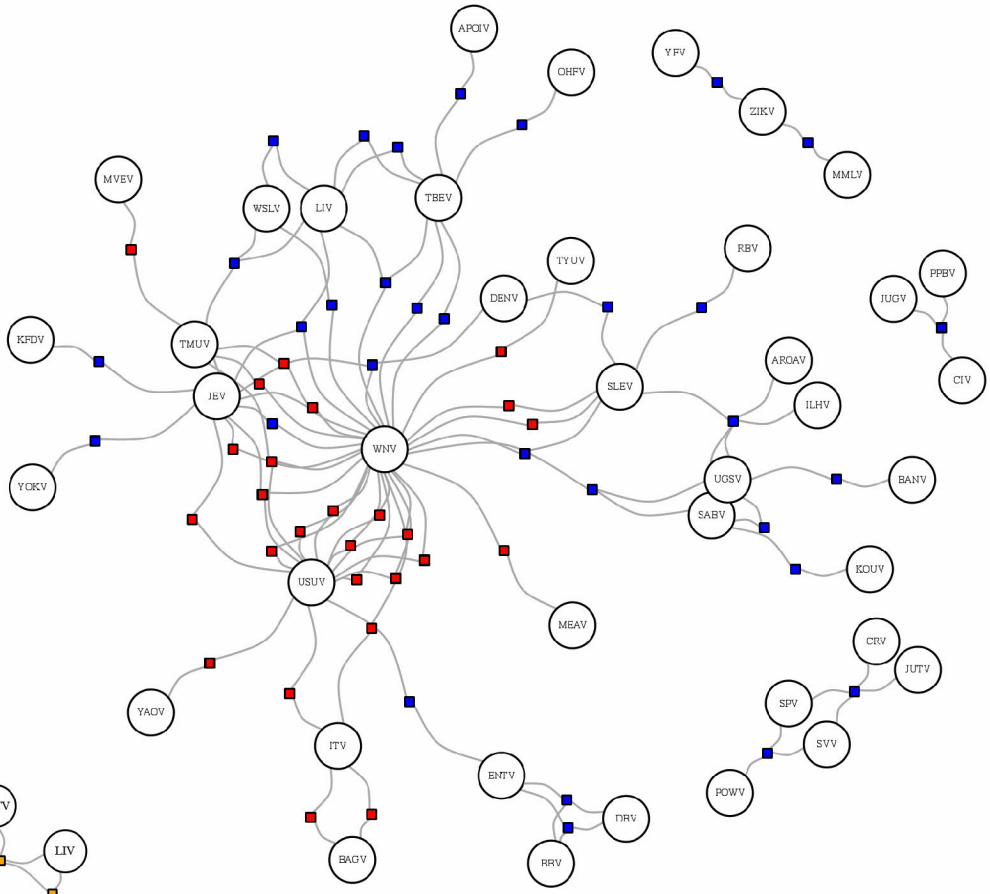
Emerging Pandemic Threats Program

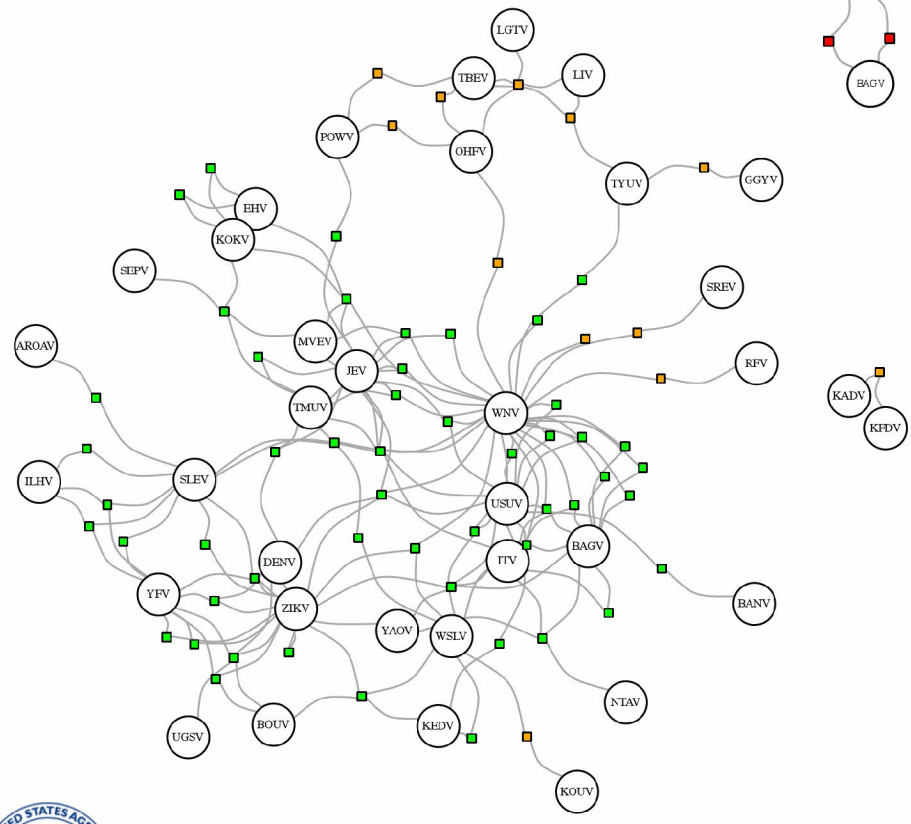
34. Global

Flavivirus Networks

Preliminary Analysis from Literature

34a. Network analysis of all the host species relationships for all known flaviviruses from the literature (right). Database includes 43 viruses (circles) detected via PCR or viral isolation in 481 hosts (squares, with mammals in blue and birds in red). 418 dead end host species are not shown.





34b. Network analysis of all the vector associations for all known flaviviruses from the literature (left). Database includes 35 viruses (circles) found in 63 vector species (squares, Insecta in green and Arachnida in orange). 97 dead end vectors are not shown.

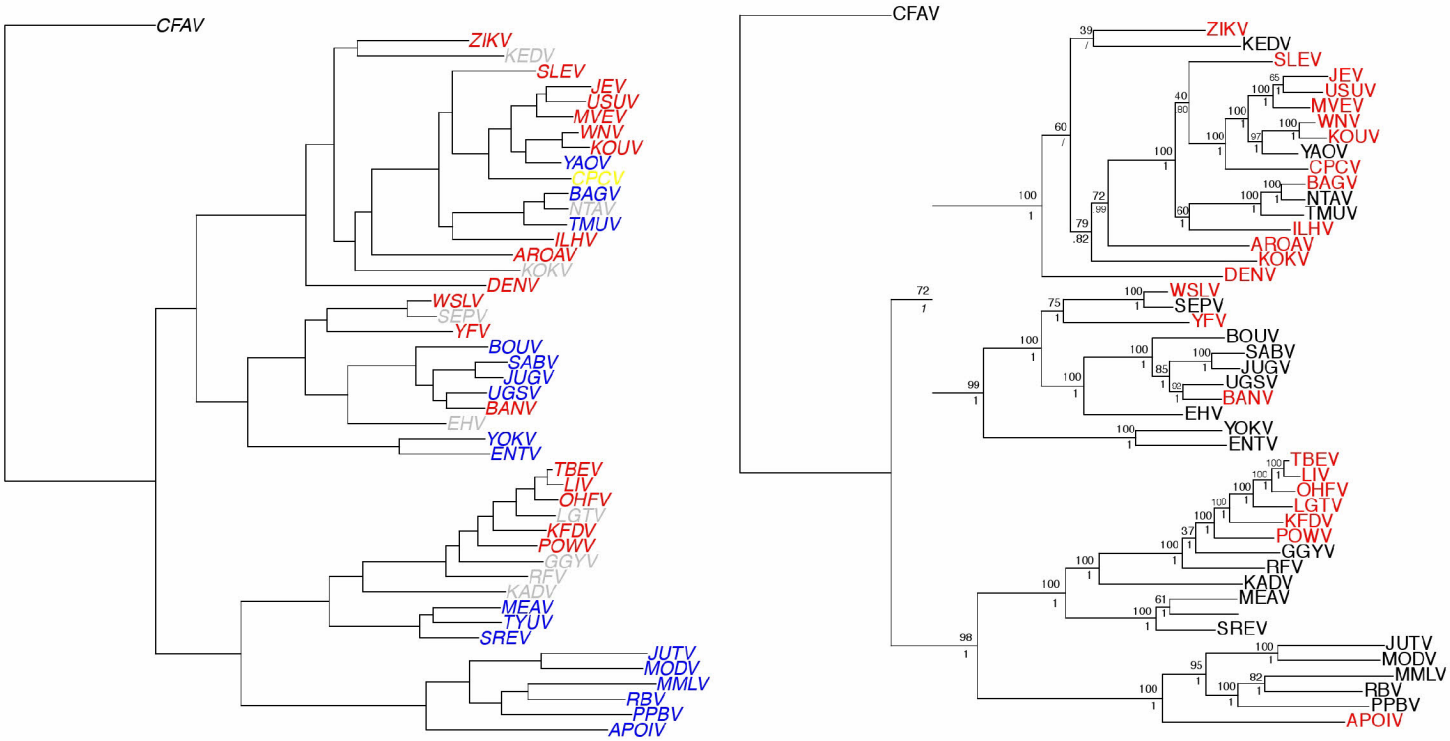
USAID
FROM THE AMERICAN PEOPLE

Emerging Pandemic Threats Program

UCDUSR0020034

35. Global

Flavivirus Phylogenetic Trait Analysis



35. Following the emergence of Zika virus, we examined what other flaviviruses might have potential to infect people. Here, we analyzed whether close relatives of zoonotic flaviviruses also tended to be zoonotic. The results can be used to deduce which of these has potential to emerge in people.

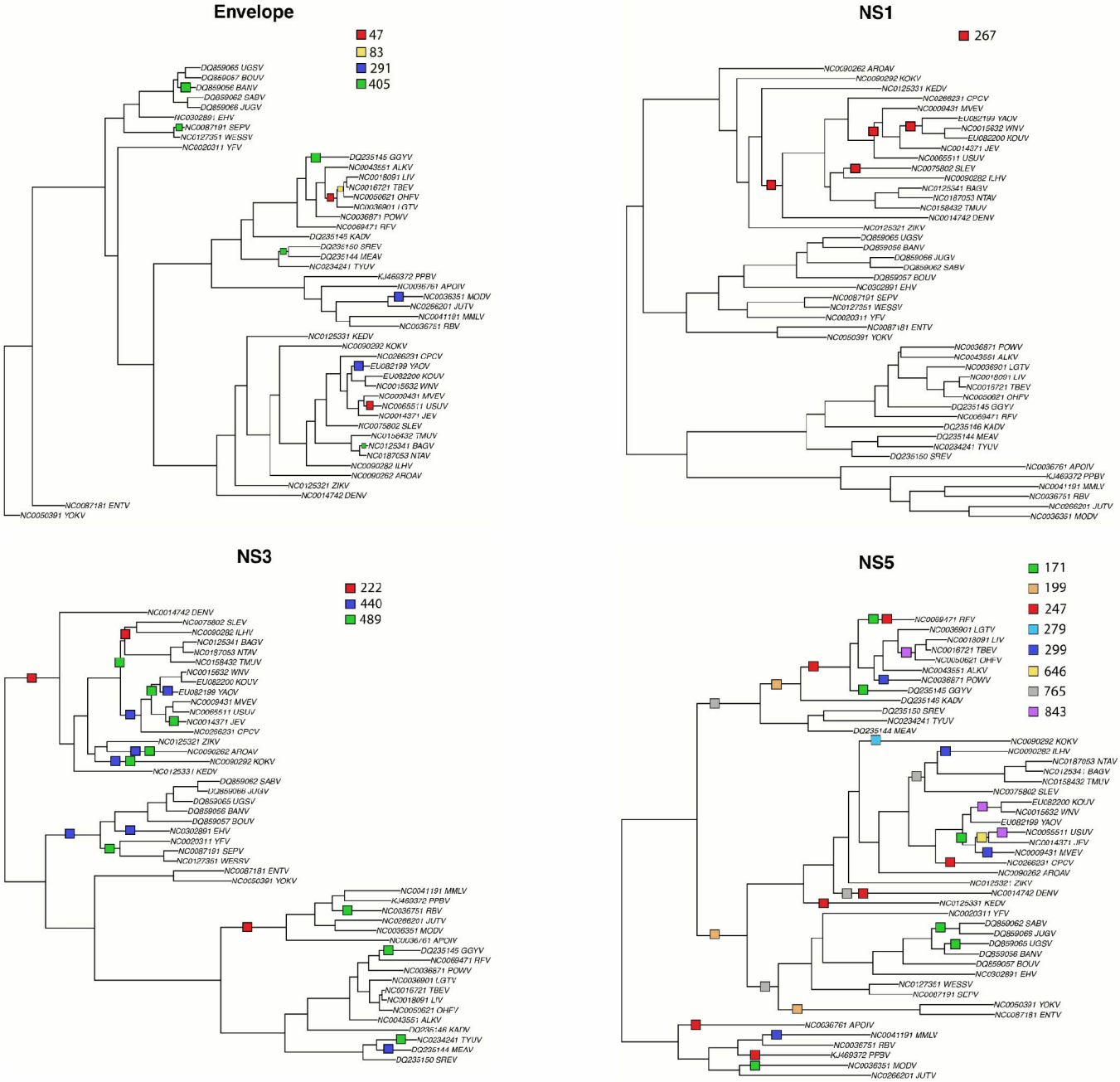
Viral phylogenetic trees showing: a) whether or not a virus is zoonotic; and b) whether or not viruses causes clinical infection in humans. Consensus phylogeny constructed with 46 full genome GenBank type species and analyzed with Bayesian tip-siginificance testing (BaTS) program. Viruses with human clinical symptoms are more likely to cluster together in the E, NS1, NS5 and consensus gene phylogenies (p = 0.001).

USAID FROM THE AMERICAN PEOPLE

Emerging Pandemic Threats Program

36. Global

Flavivirus Amino Acid Selection Analysis



36. Analysis to identify specific genetic changes that may be associated with a Flavivirus being zoonotic.

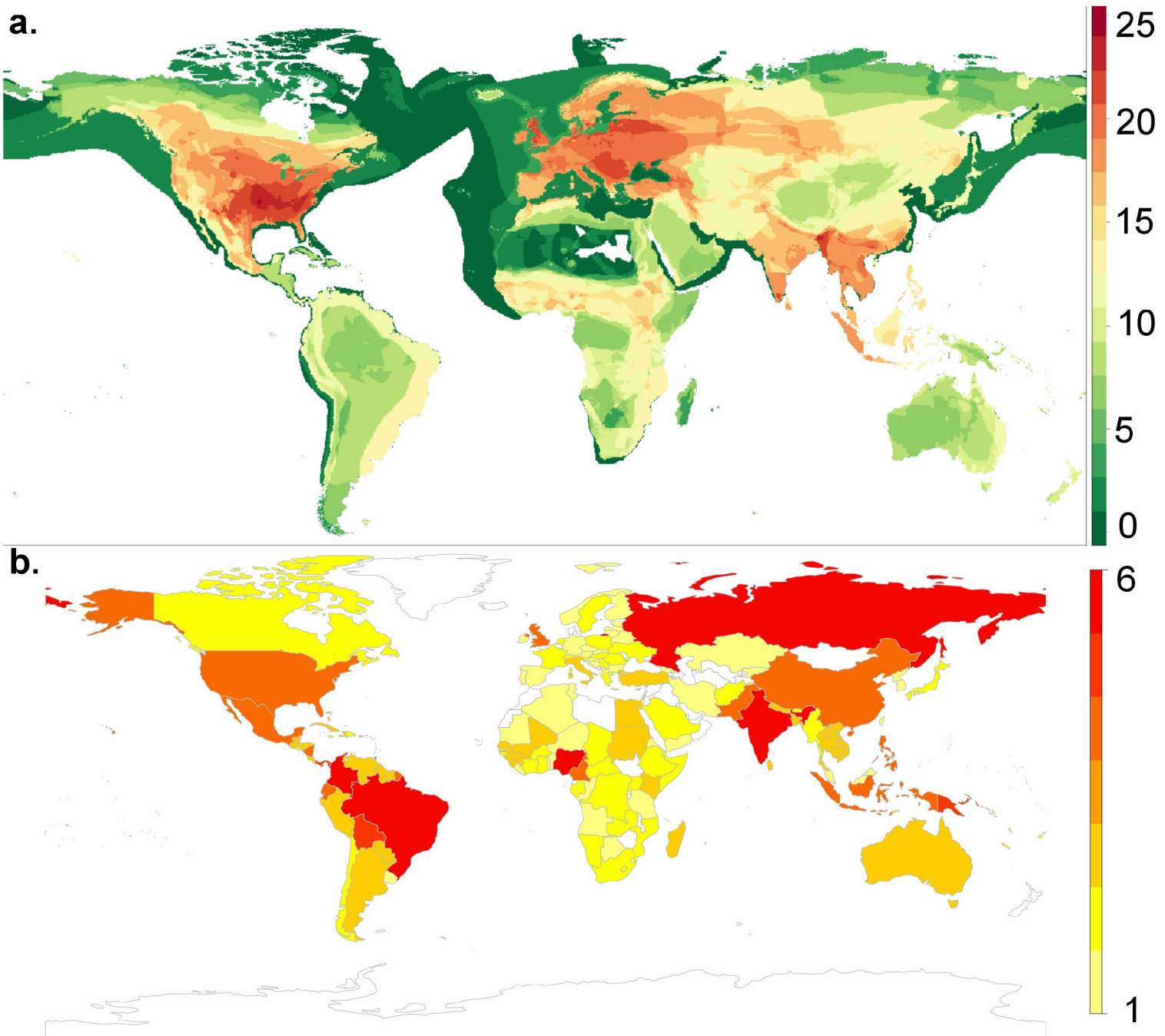
Specific amino acids under selection were identified using a full genome analysis of 46 Flaviviruses. Genes with most coverage (Envelope, NS1, NS3, and NS5) are shown. 16 codons (colored squares) were found to be positively selected through at least one of five models.

USAID FROM THE AMERICAN PEOPLE

Emerging Pandemic Threats Program

37a-b. Global

Flavivirus Host Geographic Range



37. a) Map showing where the most unique clusters of flaviviruses are on the planet. b) Map shows the diversity of 17 Flaviviruses known to cause human clinical infection in 187 countries (extracted from GIDEON). Redder regions show higher concentration of unique Flaviviruses in both maps.

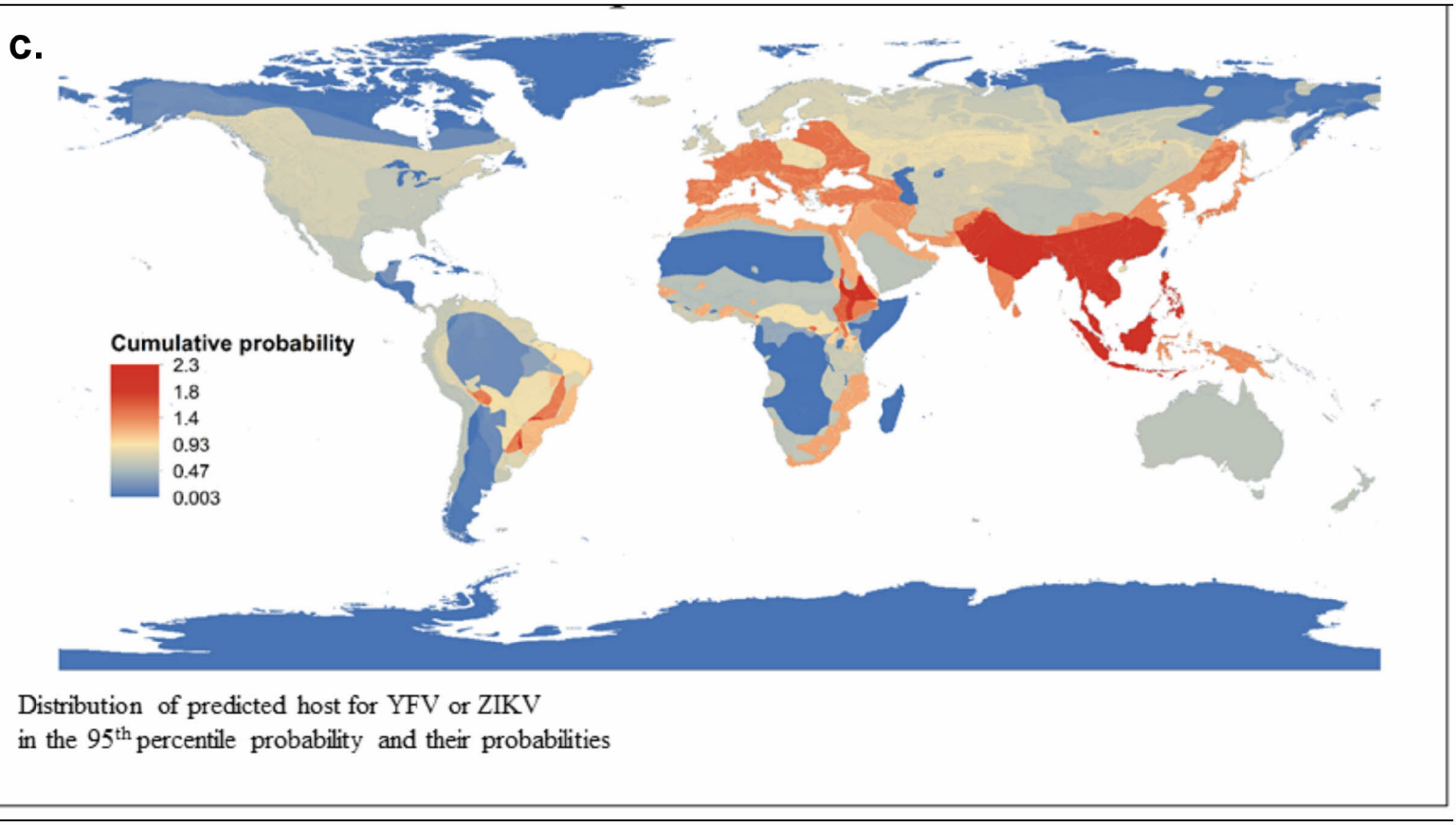
USAID
FROM THE AMERICAN PEOPLE

Emerging Pandemic Threats Program

UCDUSR0020037

37c. Global

Flavivirus Host Geographic Range



Distribution of predicted host for YFV or ZIKV in the 95th percentile probability and their probabilities

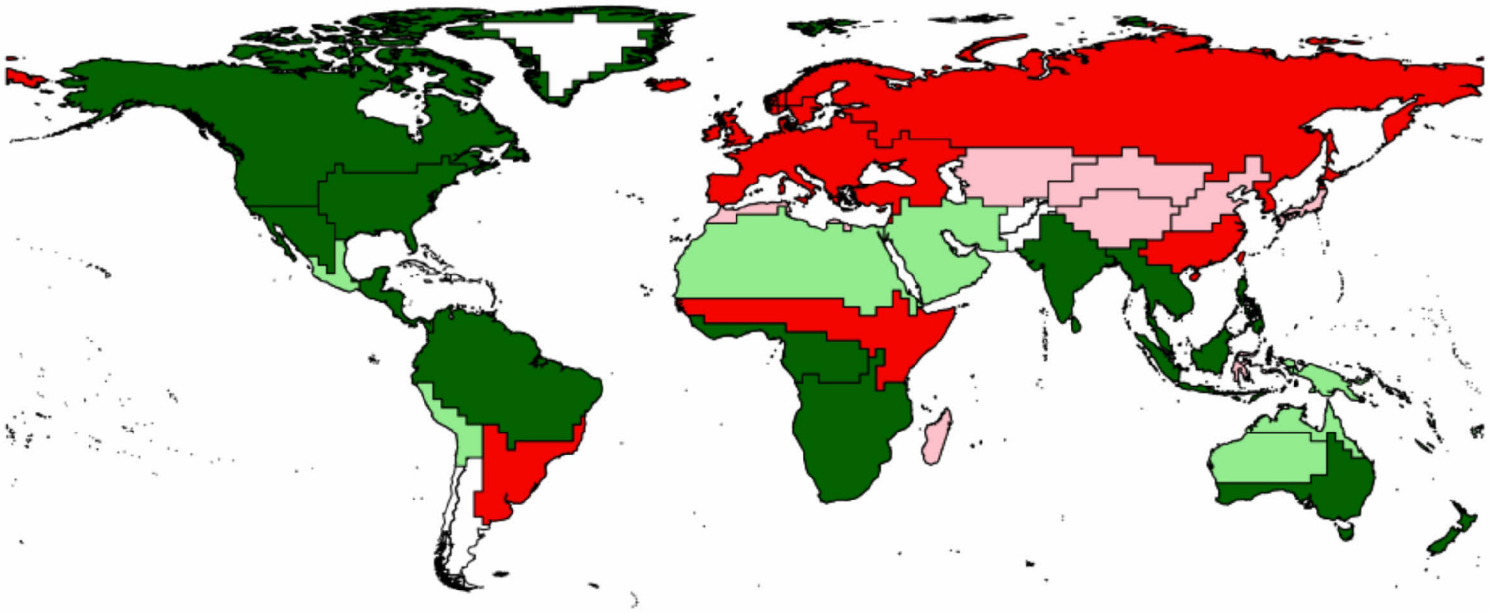
37. c) Characterizing wildlife hosts of zoonotic flaviviruses to predict zoonotic reservoir hosts and inform surveillance activities. This modelling activity seeks to build on PREDICT surveillance efforts by identifying likely sylvatic hosts of flaviviruses beyond the wild animal hosts discovered to date. Data have been collected on all known and suspected flavivirus hosts and vectors, and we used a gradient boosted regression tree model to predict hosts with a high probability of serving as a host for different flaviviruses, based on their ecological traits, behaviour, and interactions with known flavivirus vectors. This figure shows the distribution of predicted hosts for Zika virus and Yellow fever virus and their probabilities of being a host for one of the viruses. Subsequent modelling predictions seek to further evaluate differences in flavivirus predilictions for different hosts and geographic areas with wildlife hosts that have a high probability of enabling sylvatic cycles of flavivirus transmission that can potentiate viral spread and continued spillover to human populations.

USAID FROM THE AMERICAN PEOPLE

Emerging Pandemic Threats Program

38. Global

Spatial Bias in Host-Specific Zoonoses Models



38. Designed new approach to predict the number of zoonotic viruses per mammalian host species. Map shows where the most reliable data are, and where biases reduce the quality of predictions.

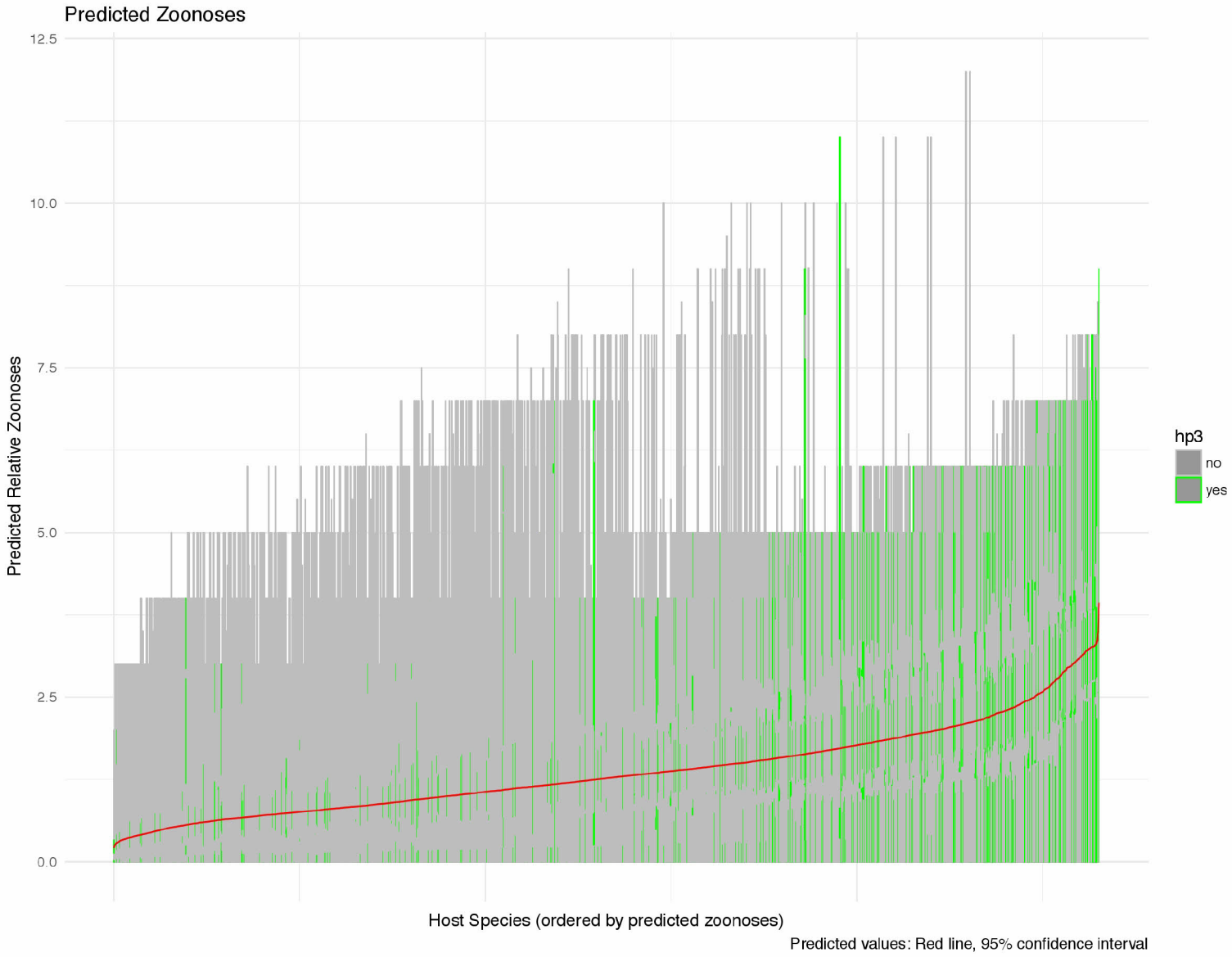
Dark green areas on the map indicate unbiased regions, while dark red indicates regions with evidence of geographically biased predictions. Lighter colors indicate regions with less than ten assigned species in the analysis. The code developed for this model has been shared with the public to allow improvements in evaluating subsequent predictive modeling efforts by other researchers.

USAID
FROM THE AMERICAN PEOPLE

Emerging Pandemic Threats Program

39a. Global

Predicting zoonotic viral richness for all mammals



39. Expanded a global model to predict the number of zoonotic viruses per mammalian host species.

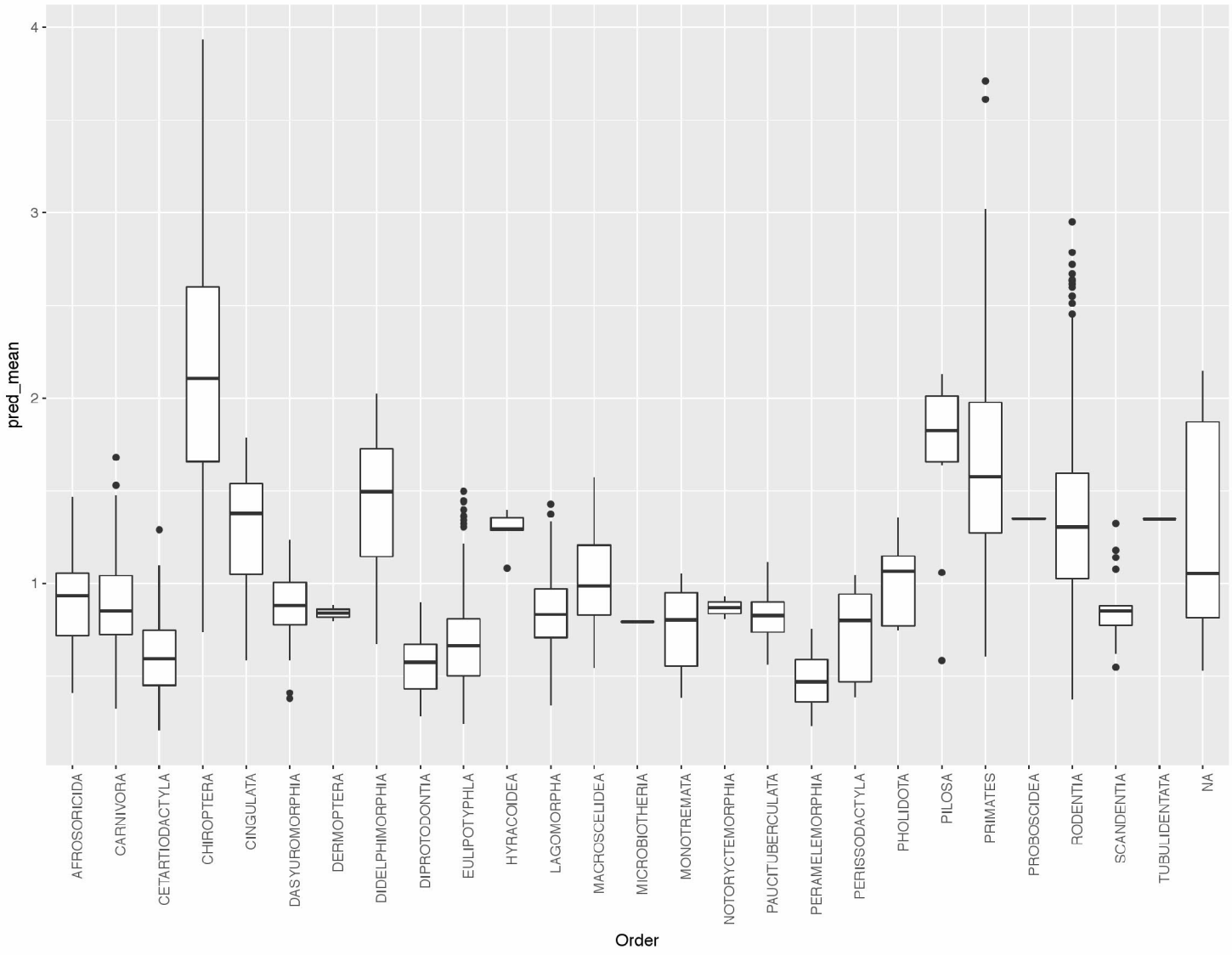
Analysis takes data from all mammals known to harbor at least one known virus, and extends it to all 5304 terrestrial mammalian hosts. Plot shows the relative number of predicted zoonotic viruses per species, ordered by value. Green area show the 656 species included in our previous missing zoonoses analysis, grey show 4648 new species with estimated numbers of zoonoses.

USAID
FROM THE AMERICAN PEOPLE

Emerging Pandemic Threats Program

39b. Global

Predicting zoonotic viral richness for all mammals



39b. Analyzed data on all mammal hosts of known viruses to predict the number of zoonotic viruses they harbor. This analysis confirms the importance of PREDICT's taxonomic focus in bats (Chiroptera), primates (Primates), and rodents (Rodentia), and also that there may be other mammalian orders of interest for maximizing zoonotic viral discovery globally.

Boxplot shows the estimated number of zoonoses per species, aggregated by order.

USAID FROM THE AMERICAN PEOPLE

Emerging Pandemic Threats Program

40. Global Viral Ranking Expert Survey

VIRAL RISK RANKING PARTICIPANT FORM

Age	
Gender	
Country of Residence	
Employer	
Academic qualification(s) and subject	
Primary field of expertise	
Additional field of expertise	
Number of years in primary field	

CATEGORY	*CONTRIBUTION TO THE RISK OF A NEW HUMAN VIRAL SPILLOVER OR EPIDEMIC EVENT OF ANIMAL-ORIGIN*	*EXPLANATION*	*LEVEL OF EXPERTISE*
Number of host species	MEDIUM	*The number of host species the virus is known to infect*	EXPERT
Host breadth	HIGH	*The diversity (i.e number of orders) of hosts the virus is known to infect*	EXPERT
Habitat breadth of host(s)	LOW	*The total number of habitats the host(s) occupies, as recognized by the IUCN, i.e. Forest, Savanna, Shrubland. Habitat is the place or environment where a plant or animal naturally or normally lives and grows*	PROFICIENT
Diet beadth of host(s)	LOW	*The total number of food items in the diet of the host(s), as described in PANtheria database (Jones et al. 2009)*	PROFICIENT
IUCN conservation status of host(s)	MEDIUM	*The conservation status of a species is an indicator of how likely it is to remain alive at present or in the near future*	PROFICIENT
Mass of host(s)	MEDIUM		PROFICIENT

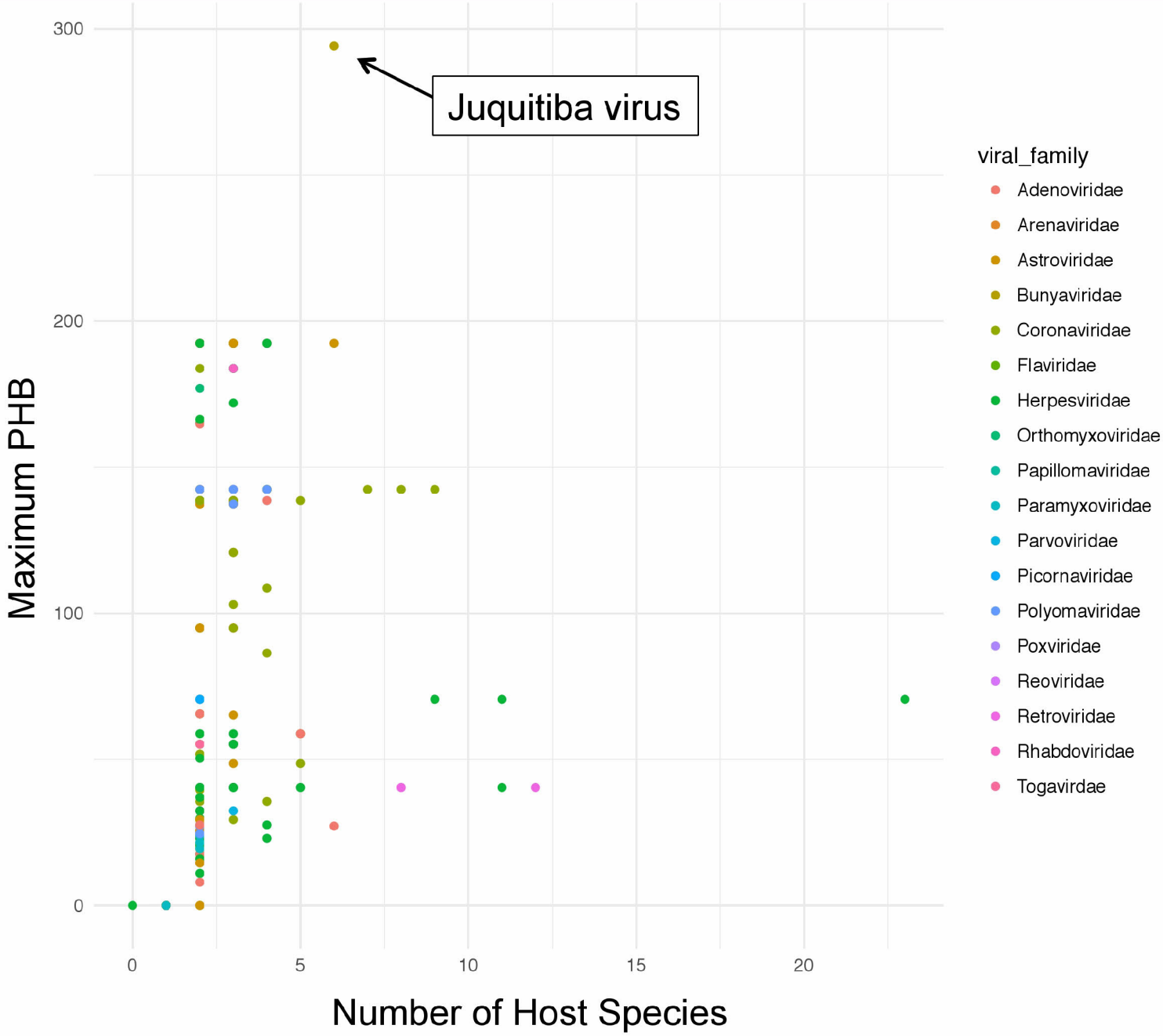
40. To assist in assessing how likely novel viruses discovered by PREDICT are to be able to infect people, we ran a survey among 150 public health professionals, virologists, and disease researchers. The ranking system will help us decide which viruses require further, in-depth characterization and study to prevent their emergence.

USAID FROM THE AMERICAN PEOPLE

Emerging Pandemic Threats Program

41. Global

PREDICT-1 Viral Phylogenetic Host Breadth



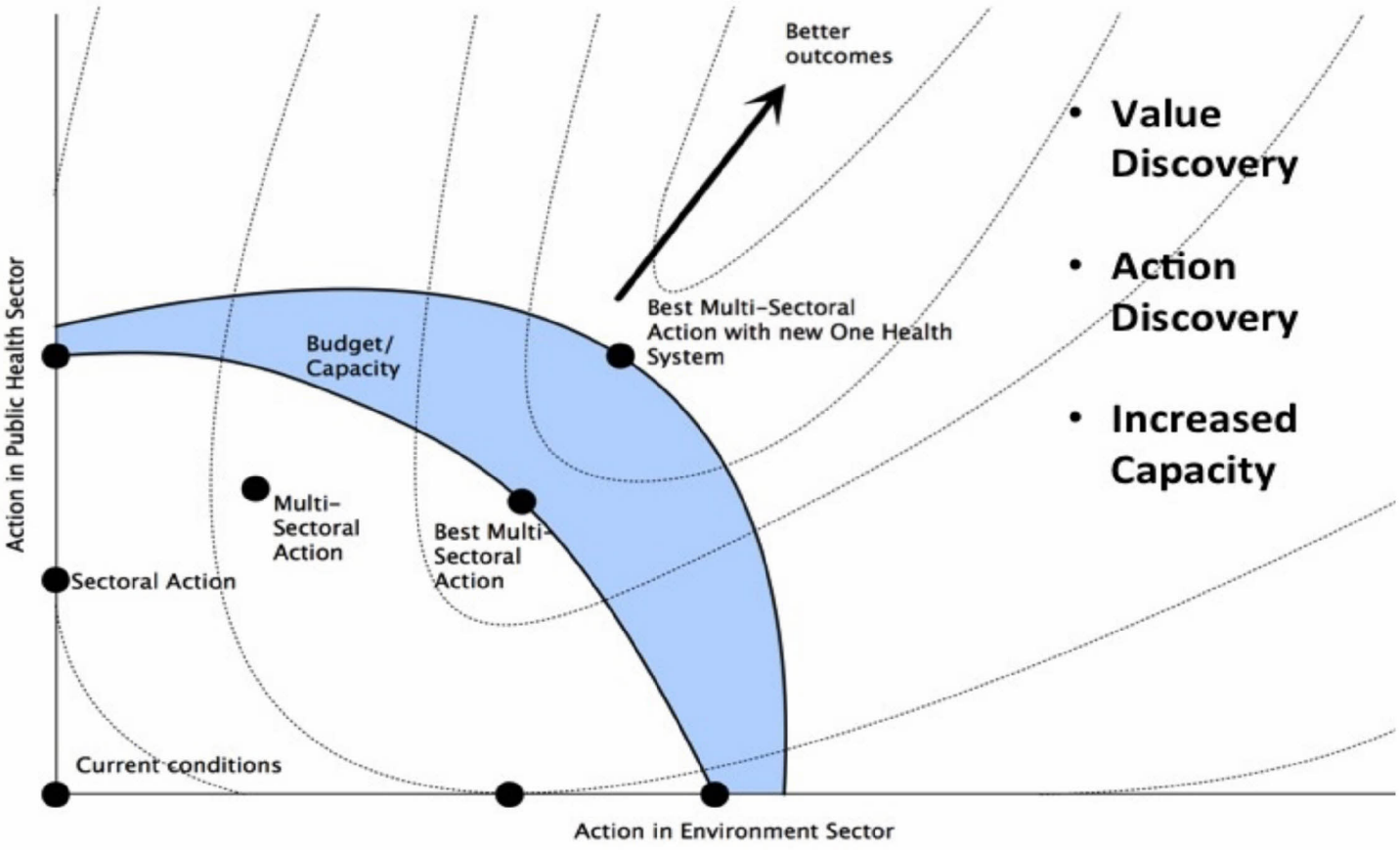
41. Developed a metric to help rank novel viruses by their zoonotic potential. Calculated the "Phylogenetic Host Breadth" (genetic diversity of all known hosts for a virus) for all P1 viruses based our surveillance data of host species found positive. The virus with greatest PHB from PREDICT data is Juquitiba virus (highlighted). This typically rodent-borne Hantavirus was detected by PREDICT1 in bats and a opossum.

USAID FROM THE AMERICAN PEOPLE

Emerging Pandemic Threats Program

42. Global

Economics and EID

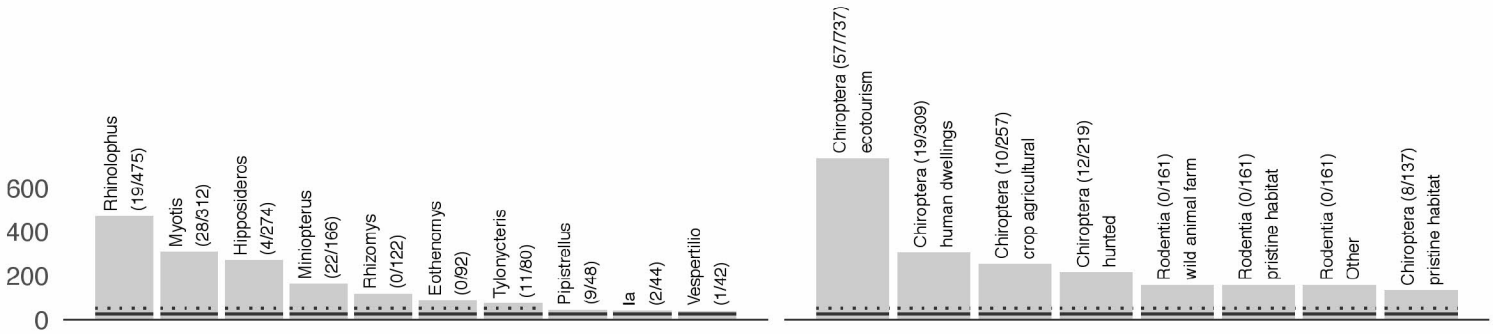


42. Developed a conceptual model to map out the impacts and opportunities of a One Health framework. While the One Health concept is typically presented as a Venn diagram with all circles equal weight, in reality the role of (and affect on) each of three sectors varies widely based on the particular context of the situation (which may be dependent on the specific disease, transmission pathway, risk factors, etc.).

This model was developed at a workshop led by PREDICT and the World Bank, entitled “Prevent, Prepare and Respond: Economics of a One Health Approach to Confront Disease Threats”.

43. Global

PREDICT-1 Power Analysis of Interfaces



Horizontal lines represent sample sizes required to detect a difference of 10% (solid: 28 animals), or 5% (dotted: 55 animals) from global baseline detection rates (~0.379126%). Parentheses show total animals with viral detections / number of animals tested. Countries ordered by total sampled animals. Only top 10 genuses sampled in each country shown.

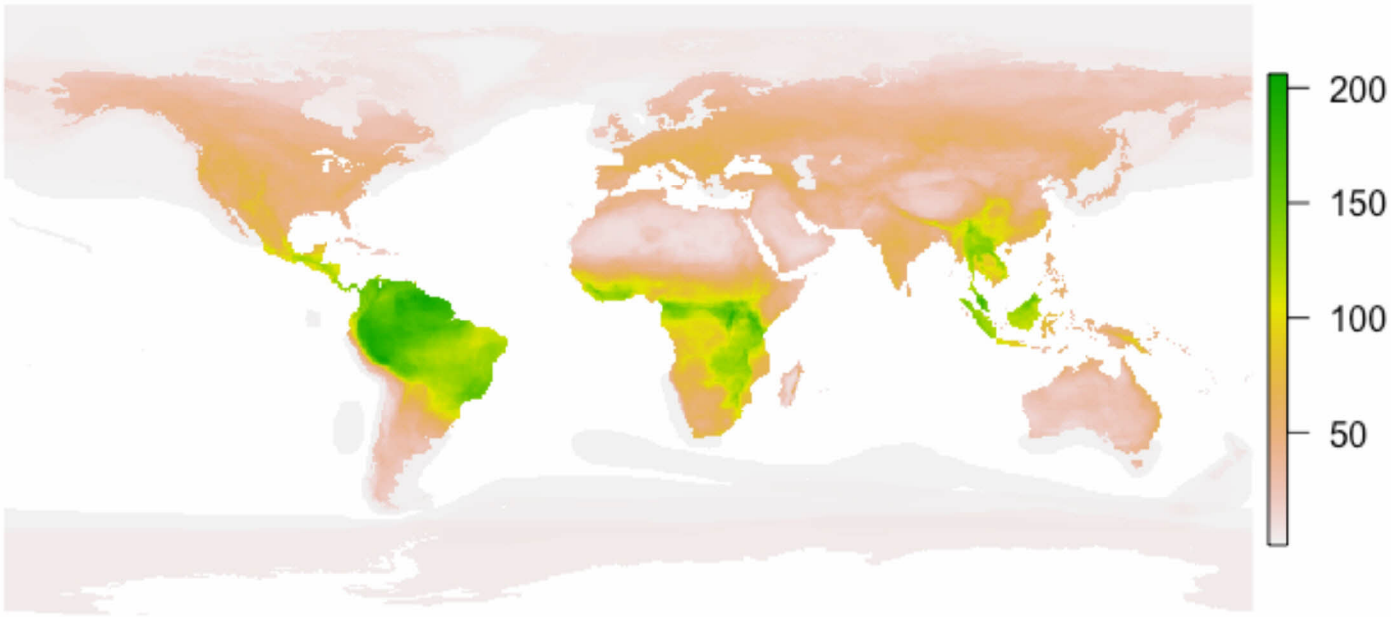
43. To determine the optimal sample sizes to discover new viruses at different risk interfaces (and taxonomic groups), we analyzed viral detection rates and animal sampling distributions from PREDICT-1 data.

USAID
FROM THE AMERICAN PEOPLE

Emerging Pandemic Threats Program

44. Global

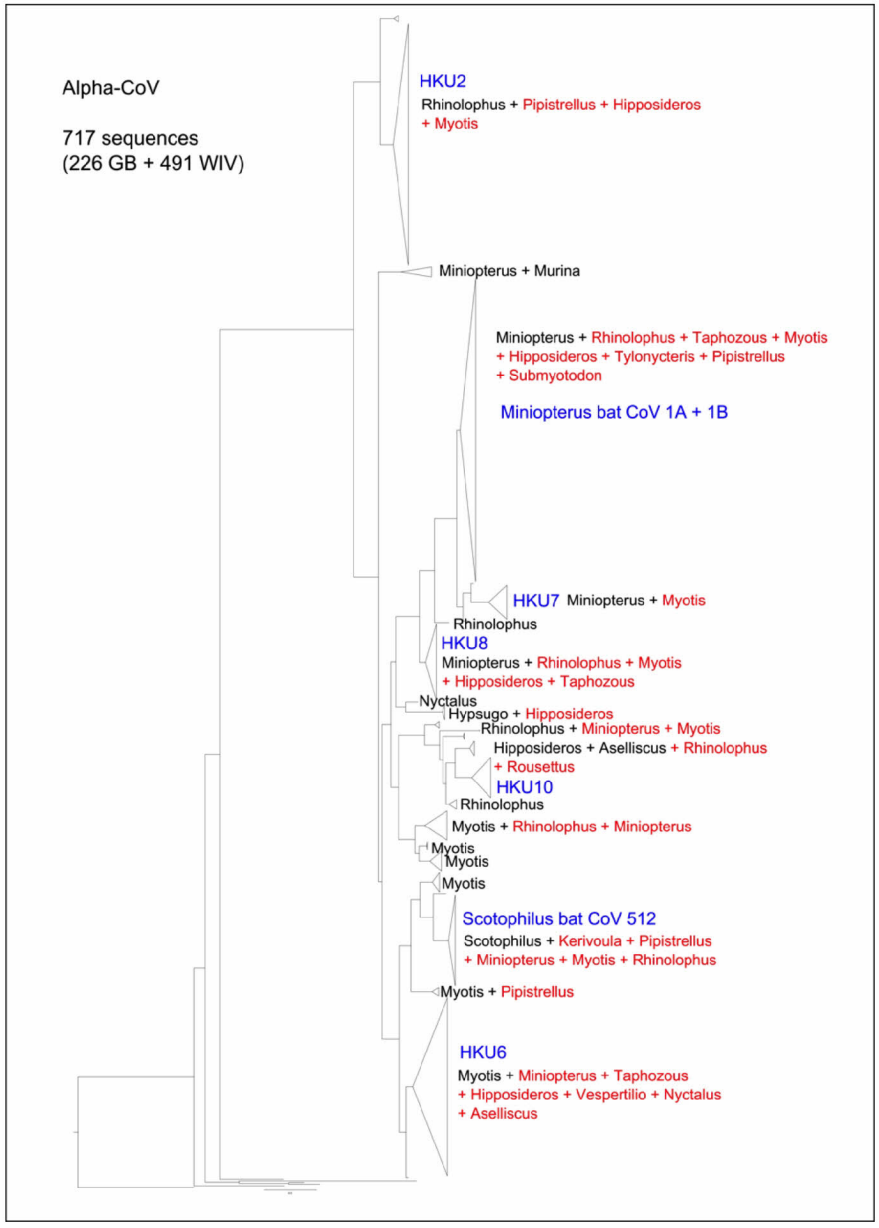
Software for Rapid Spatial Mapping



44. Developed a new software, fasterize (R package), implementing high-speed algorithms for geospatial analyses – to rapidly speed up time-consuming risk analyses. This tool will be made publically available.

45a. China

Alpha-Coronavirus Host-Virus Phylogenetic Mapping



47a. As a way to understand which PREDICT coronaviruses are more likely to jump from one host species to another (i.e. their potential to infect people) we analyzed host and virus relationships (mapped on viral phylogenies) for all coronaviruses discovered in China during PREDICT-1.

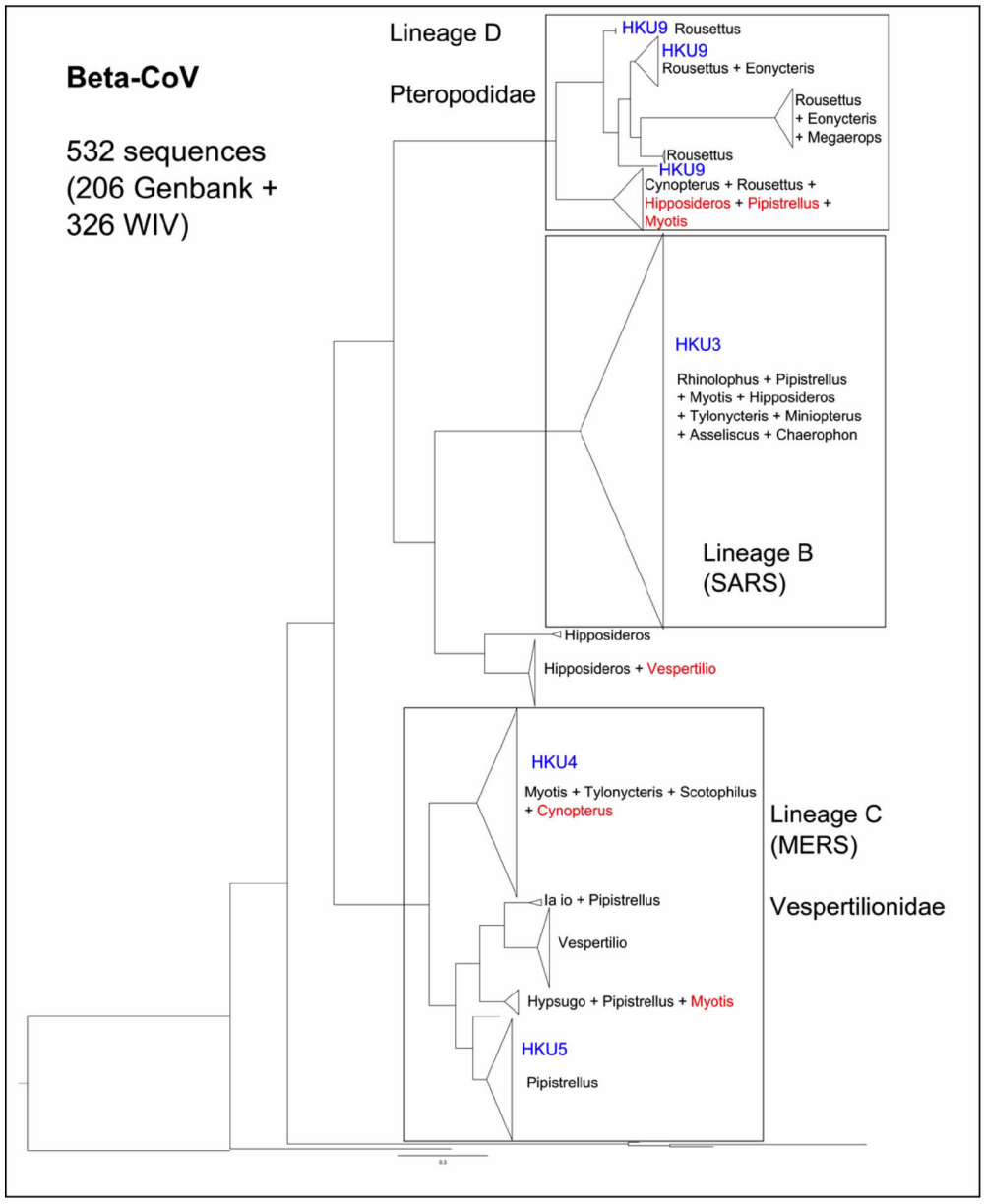
Figure shows the sub-group: alpha-coronaviruses from China. Taxa in red correspond to a very small number of sequences in the clade that may represent cross-species transmission events.

USAID FROM THE AMERICAN PEOPLE

Emerging Pandemic Threats Program

45b. China

Beta-Coronavirus Host-Virus Phylogenetic Mapping



45b. As a way to understand which PREDICT coronaviruses are more likely to jump from one host species to another (i.e. their potential to infect people) we analyzed host and virus relationships (mapped on viral phylogenies) for all coronaviruses discovered in China during PREDICT-1.

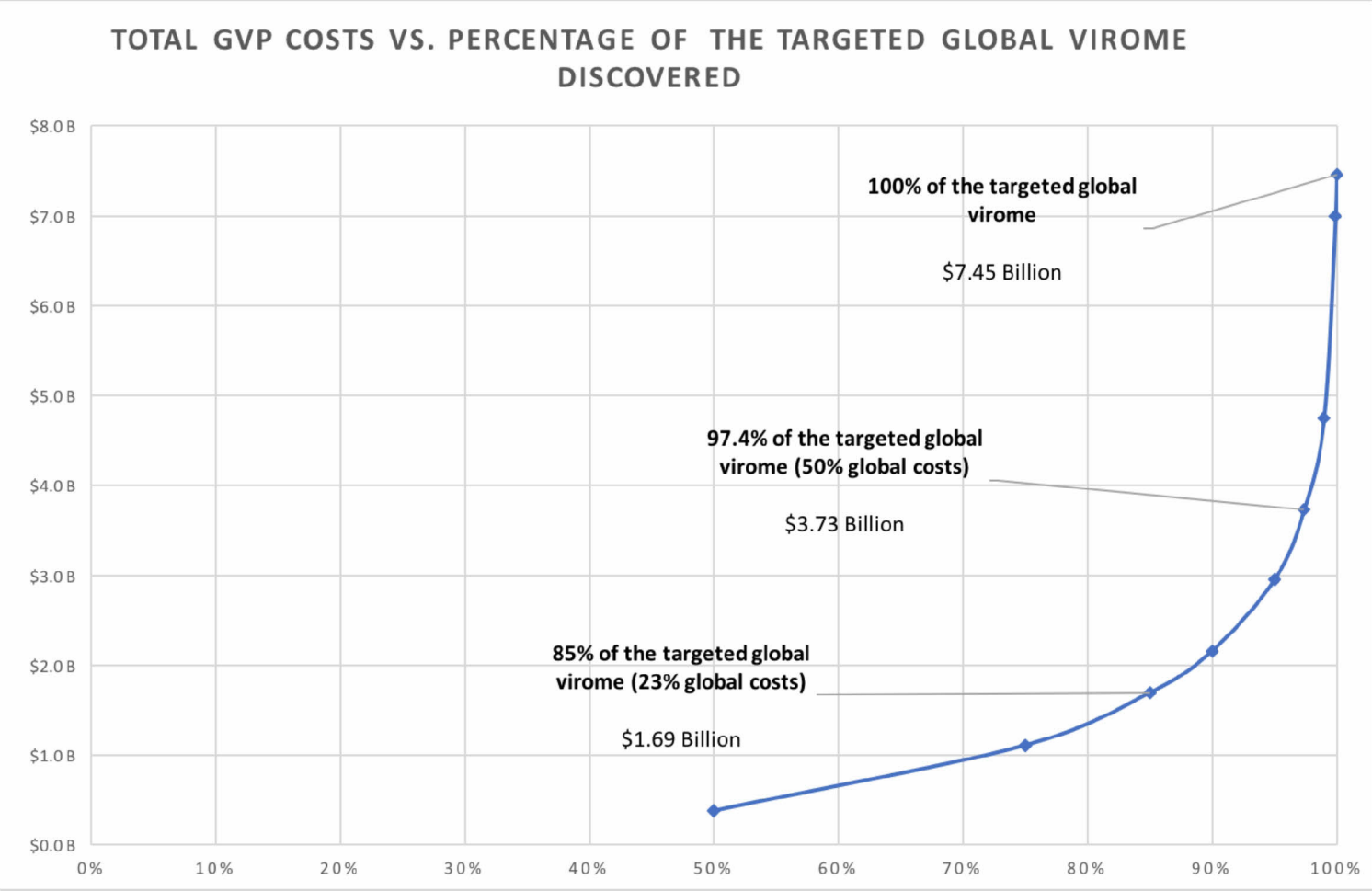
Figure shows the sub-group: beta-coronaviruses from China. Taxa in red correspond to a very small number of sequences in the clade that may represent cross-species transmission events.

USAID FROM THE AMERICAN PEOPLE

Emerging Pandemic Threats Program

46. Global

Projected Costs of the Global Virome Project



46. Projected total costs and coverage of Global Virome Project (GVP) sampling of 5,219 terrestrial mammals and 50 domestic species with testing for 24 viral families of zoonotic potential, as well as 871 species of water birds tested for influenza viruses. Cost increase exponentially to discover 100% of all viruses, as some viruses are very rare and require extremely large sample sizes to discover; the majority of viruses can be discovered (e.g. 97% or 85%) for a fraction of the cost.

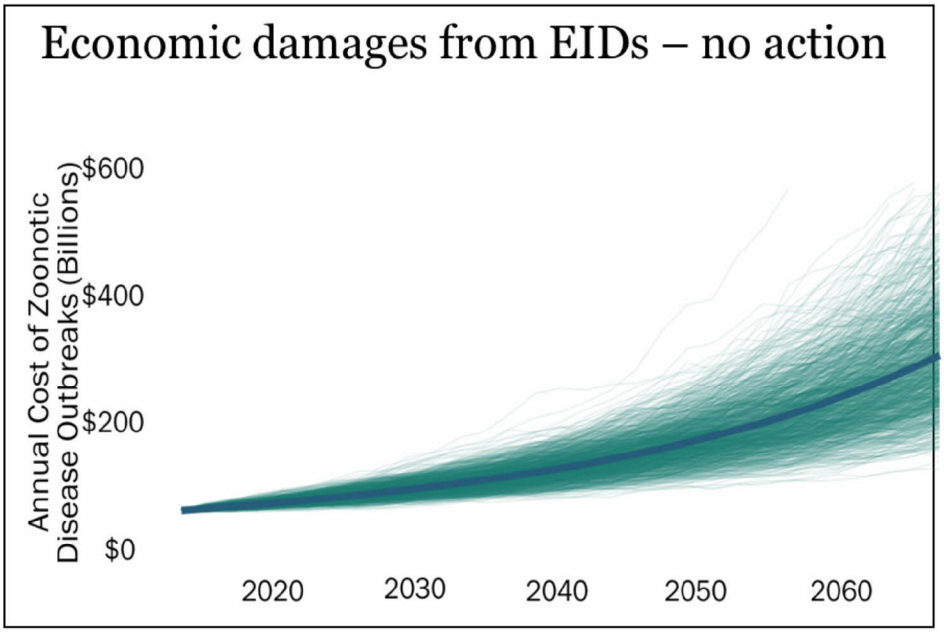
USAID FROM THE AMERICAN PEOPLE

Emerging Pandemic Threats Program

47. Global

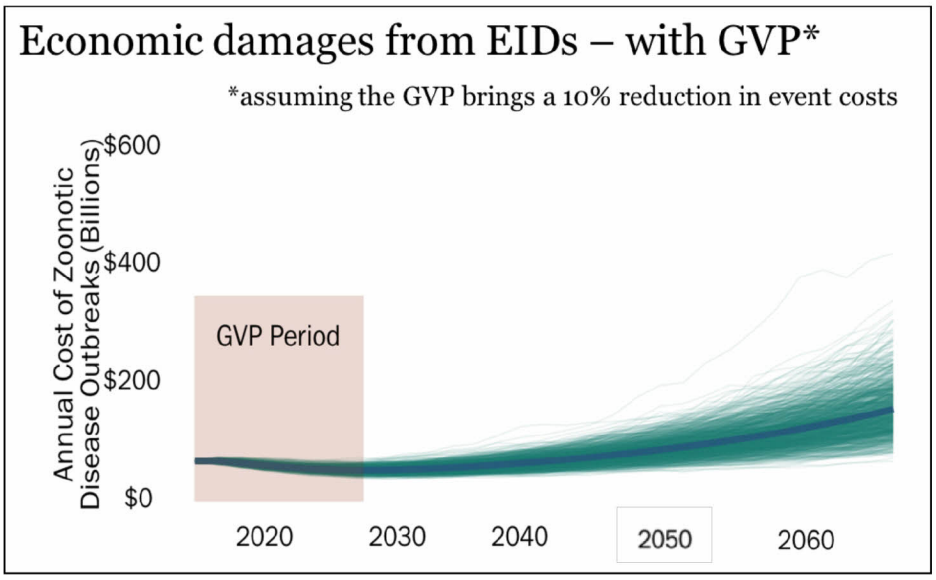
Reduction in EID Economic Damages w/ Global Virome Project

a.



Total damages:
$3.6 trillion

b.



Total savings:
$326 billion

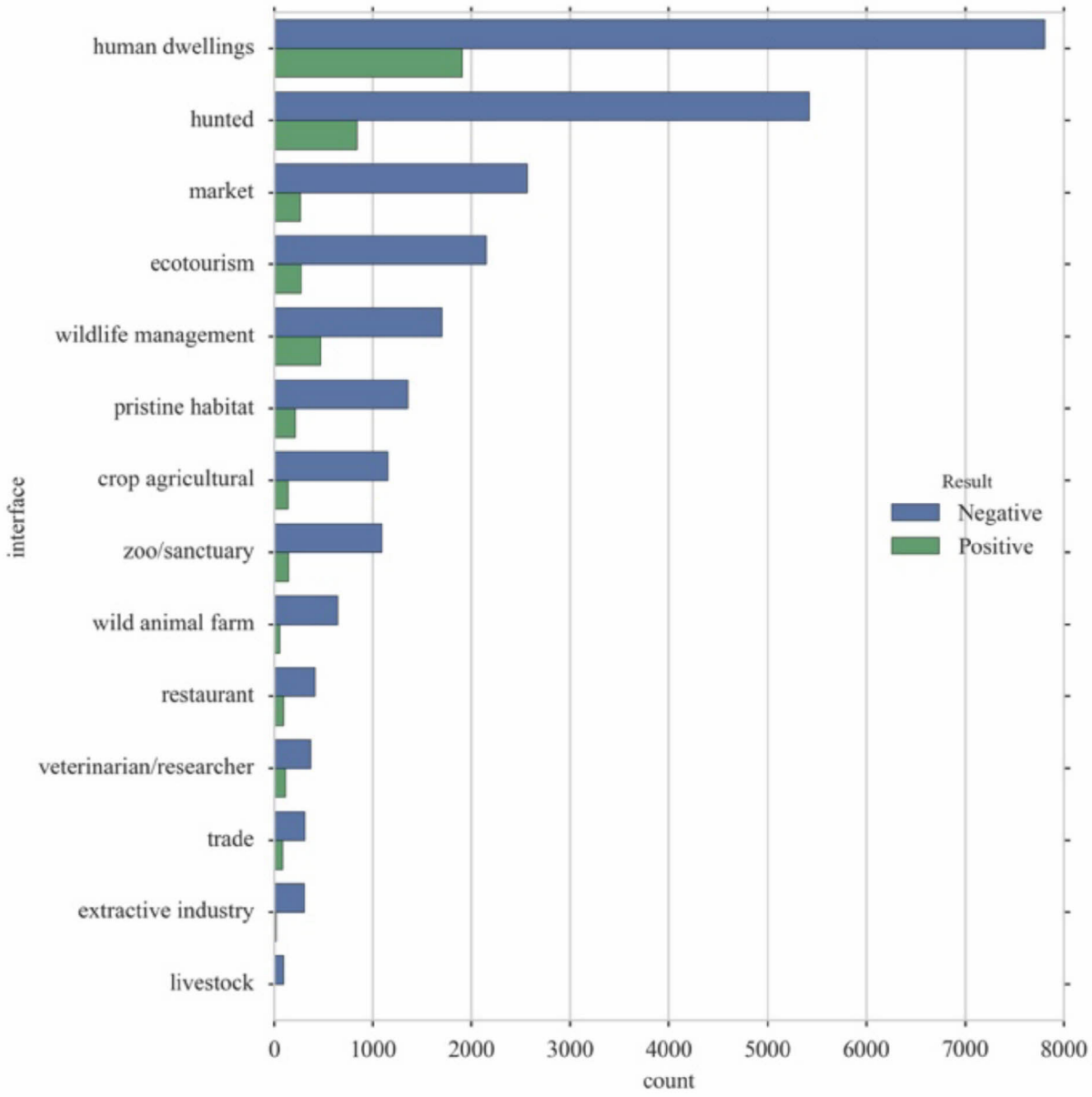
47. To assess long term trends in pandemic risk, calculated expected global economic damages from EIDs over the next 50 years using data from the Emerging Infectious Disease Repository (EIDR) – a PREDICT-derived publically-available database. A) Total cost is estimated at $3.6 trillion over the next 50 years if no coordinated global action is taken. B) Reduction in damages by investing in a control program, e.g. Global Virome Project (GVP), that reduces the number of events or size of events by only 10%.

USAID
FROM THE AMERICAN PEOPLE

Emerging Pandemic Threats Program

48a. Global

Predict-1 Viral Sharing at Animal-Human Interfaces



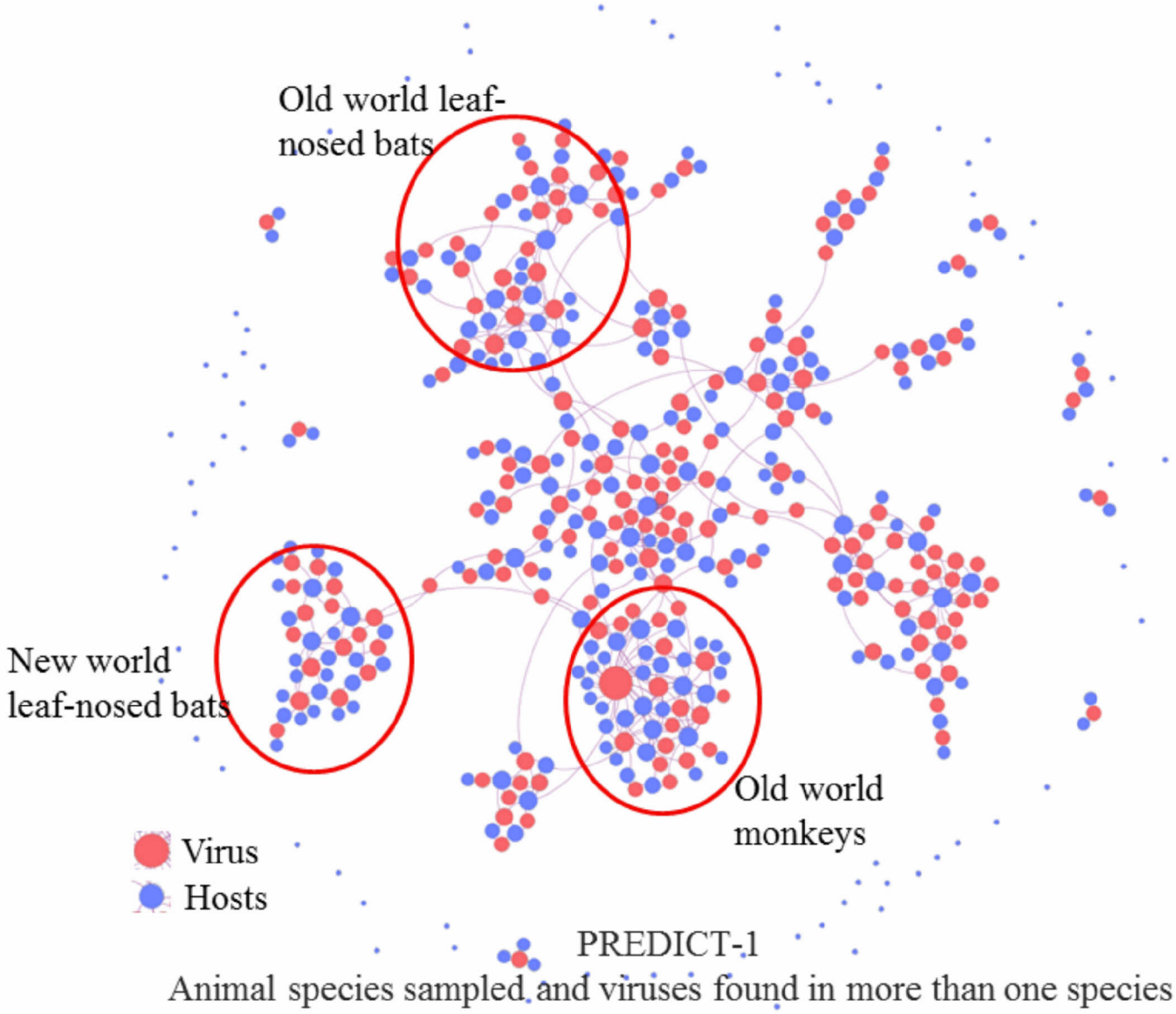
48. Risk of viral spillover was further investigated with PREDICT-1 data on the epidemiological context for human contact with animals and possible direct and indirect disease transmission routes for all animals sampled from 2009-2014. We assessed sampling and testing effort for all host-interface combinations to evaluate power to detect differences across the 17 categorized interfaces investigated in PREDICT 1. Risk of viral spillover at all wildlife-human interfaces is being further characterized based on detection of 1) known viruses, 2) novel viruses, 3) and viruses with high host plasticity, stratified by viral family as appropriate.
This figure shows the number of animals sampled at different interfaces and frequency of negative and positive test findings using PREDICT viral detection protocols (cPCR).

USAID
FROM THE AMERICAN PEOPLE

Emerging Pandemic Threats Program

48b. Global

PREDICT-1 Viral Sharing at Interfaces



48. This figure shows a bipartite host-virus network showing viruses found in more than one species. Clustering of hosts shows viruses shared in taxonomically related host species. Analyses are providing insight to optimize surveillance activities in PREDICT-2 and characterize risk for viral spillover and spread.

USAID FROM THE AMERICAN PEOPLE

Emerging Pandemic Threats Program